TASK: PA29
CDRL: AC01
31 December 1996

INFORMAL TECHNICAL REPORT

For

SOFTWARE TECHNOLOGY FOR ADAPTABLE, RELIABLE SYSTEMS
(STARS)

Canvas Knowledge Acquisition Guidebook Version 2.0

STARS-PA29-AC01/001/01

31 December 1996

CONTRACT NO. F19628-93-C-0130

Prepared for:
Electronic Systems Center
Air Force Systems Command, USAF
Hanscom, AFB, MA 01731-2816

Prepared by:
Lockheed Martin C^2 Integration Systems
9255 Wellington Road
Manassas, VA 20110-4121

TASK: PA29
CDRL: AC01
31 December 1996

Data Reference: STARS-PA29-AC01/001/01
INFORMAL TECHNICAL REPORT
Canvas Knowledge Acquisition Guidebook
Version 2.0

TASK: PA29
CDRL: AC01
31 December 1996

Data Reference: STARS-PA29-AC01/001/01
INFORMAL TECHNICAL REPORT
Canvas Knowledge Acquisition Guidebook
Version 2.0

Abstract

This guidebook describes the Canvas approach to systematic knowledge acquisition. Canvas synthesizes elements of two distinct methods: Scenario-Based Engineering Process (SEP) and Organization Domain Modeling (ODM). SEP provides knowledge acquisition methods and representation techniques. ODM provides a conceptual framework for data acquisition planning for the purposes of domain engineering. The guidebook incorporates extensive lessons learned from project experience in managing a large-scale knowledge acquisition effort in coordination with advanced technology development in the health-care domain.

TASK: PA29
CDRL: AC01
31 December 1996

Data Reference: STARS-PA29-AC01/001/01
INFORMAL TECHNICAL REPORT
Canvas Knowledge Acquisition Guidebook
Version 2.0

Principal Author(s):

Mark Simos, Organon Motives, Inc.	Date
Dean Allemang, Organon Motives, Inc.	Date
Charles Hammons, ScenPro, Inc.	Date
Lisa Mantock, ScenPro, Inc.	Date
Carol Klingler, Lockheed Martin C^2 Integration Systems	Date
Larry Levine, Organon Motives, Inc.	Date
Dick Creps, Lockheed Martin C^2 Integration Systems	Date

Approvals:

Teri F. Payton, Lockheed Martin C2 Integration Systems STARS Program Manager	Date

(Signatures on File)

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave Blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	31 December 1996	Informal Technical Data

4. TITLE AND SUBTITLE

Canvas Knowledge Acquisition Guidebook, Version 2.0

5. FUNDING NUMBERS

F19628-93-C-0130

6. AUTHOR(S)

Mark Simos, Dean Allemang, Larry Levine, Organon Motives; Charles Hammons, Lisa Mantock, ScenPro, Inc.; Carol Klingler, Dick Creps, Lockheed Martin C2 Integration Systems.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Lockheed Martin C^2 Integration Systems
9255 Wellington Road
Manassas, VA 20110-4121

8. PERFORMING ORGANIZATION REPORT NUMBER

Document Number
STARS-PA29-AC01/001/01

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Department of the Air Force
ESC/ENS
Hanscom AFB, MA 01731-2816

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

AC01

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Distribution "A"

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This guidebook describes the Canvas approach to systematic knowledge acquisition. Canvas synthesizes elements of two distinct methods: Scenario-Based Engineering Process (SEP) and Organization Domain Modeling (ODM). SEP provides knowledge acquisition methods and representation techniques. ODM provides a conceptual framework for data acquisition planning for the purposes of domain engineering. The guidebook incorporates extensive lessons learned from project experience in managing a large-scale knowledge acquisition effort in coordination with advanced technology development in the health-care domain.

14. SUBJECT TERMS

15. NUMBER OF PAGES

184

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	SAR

Data Reference: STARS-PA29-AC01/001/01
INFORMAL TECHNICAL REPORT
Canvas Knowledge Acquisition Guidebook
Version 2.0

Table of Contents

Table of Contents

Data Reference: STARS-PA29-AC01/001/01
INFORMAL TECHNICAL REPORT
Canvas Knowledge Acquisition Guidebook
Version 2.0

Table of Contents

Data Reference: STARS-PA29-AC01/001/01
INFORMAL TECHNICAL REPORT
Canvas Knowledge Acquisition Guidebook
Version 2.0

Table of Contents

Data Reference: STARS-PA29-AC01/001/01
INFORMAL TECHNICAL REPORT
Canvas Knowledge Acquisition Guidebook
Version 2.0

List of Exhibits

Prologue

This document is version 2.0 of the *Canvas Knowledge Acquisition Guidebook.* It proposes a new and innovative approach to planning and managing knowledge acquisition activities in engineering projects. The Canvas approach synthesizes concepts and techniques from:

- The Organization Domain Modeling (ODM) domain engineering method sponsored by the DARPA Software Technology for Adaptable, Reliable Systems (STARS) program, and
- The Scenario-based Engineering Process (SEP) method sponsored by the DARPA Healthcare Information Infrastructure (HIIP) program.

Canvas was developed collaboratively by the STARS and HIIP programs under the DARPA Noah's Ark initiative (see Section 2 for further background).

Version 2.0 incorporates additional lessons learned from a brief trial application and additional experience gained in applying and teaching Canvas techniques. The most extensive changes are in Section 4.0, which now provides a more prescriptive framework for planning a knowledge acquisition effort. We have also tried to integrate the various core concepts into a more cohesive whole.

We intend to apply and evolve the Canvas approach in the future. We encourage trial use of Canvas and solicit reader review and comments as input to its future evolution. To learn more about Canvas, discuss how to obtain support in applying it, or submit feedback on the guidebook or your practical Canvas experiences, please contact one or more of the following people:

Mark Simos
Organon Motives, Inc.
One Williston Road, Suite 4
Belmont, MA 02178
Phone: (617) 484-3383 x11
Fax: (617) 383-3363
E-mail: mas@organon.com

Charles "Bud" Hammons
ScenPro, Inc.
2604 Spring Lake Drive
Richardson, TX 75082
Phone: (214) 307-7641
Fax: (214) 306-6818
E-mail: hammons@onramp.net

Dean Allemang
Organon Motives, Inc.
One Williston Road, Suite 4
Belmont, MA 02178
Phone: (617) 484-3383 x13
Fax: (617) 383-3363
E-mail: dta@organon.com

Dick Creps
Lockheed Martin
9255 Wellington Road
Manassas, VA 20110
Phone: (703) 367-1353
Fax: (703) 367-1389
E-mail: dick.creps@lmco.com

Thank you for your interest.

1.0 Document Overview

This guidebook describes an approach to systematic knowledge acquisition called *Canvas*. Canvas synthesizes elements of two distinct methods: the Scenario-based Engineering Process (SEP) and Organization Domain Modeling (ODM). SEP provides knowledge acquisition methods and representation techniques, while ODM provides a conceptual framework for data acquisition planning for the purposes of domain engineering. The guidebook incorporates extensive lessons learned from particular project experience in managing large-scale knowledge acquisition efforts, and generalizes them to a method for managing knowledge acquisition that is compatible with the goals of both ODM and SEP.

At the heart of the Canvas approach is a framework for identifying and managing the difficult issues of the knowledge acquisition process. These issues include managing the interests of the various stakeholders to the knowledge acquisition enterprise (e.g., the people providing domain knowledge, the knowledge acquisition staff, and the targeted audience for the information gathered), managing access to written sources of information, and managing the record of the knowledge that has been acquired. The Canvas framework suggests more systematic ways of planning and flexibly coordinating the knowledge acquisition effort, to anticipate and handle the influence of bias and potential sources of resistance, avoid inefficient use of experts' time, and respond to shifting availability of resources and newly discovered information sources.

The remainder of this overview section describes the document's purpose and scope, intended audience, and organization.

1.1 Purpose and Scope

Canvas is primarily concerned with planning knowledge acquisition activities within the context of a broader project such as a technology development effort. In some types of projects, such as expert systems development projects or domain engineering efforts to support systematic software reuse, knowledge acquisition is recognized as a distinct project component. In other types of system development efforts, knowledge acquisition activities may take place intermixed with other concerns such as requirements gathering; and so important dynamics of these activities may not be fully acknowledged by planners. (Section 3.1 discusses the kinds of activities that qualify as KA from the Canvas perspective, and differentiates knowledge acquisition from related types of knowledge transfer or learning activities.)

There are a wide variety of materials available for specific knowledge acquisition techniques (e.g., interviewing, data analysis, observation techniques) for many disciplines. However, planning aspects, particularly for large-scale efforts, have been much less thoroughly covered.The primary purpose of this document is to clarify the value of explicitly planning the knowledge acquisition aspects of a broad range of technology development efforts, and to provide guidelines useful in that planning process. The documented approach should be of use in other contexts as well, such as data gathering for social scientific research purposes, but will be most relevant where the presence of technological systems is a significant factor in the environments studied.

The emphasis on "planning" may make it seem as though Canvas does not cover knowledge acquisition activities proper. In fact, the planning framework is a useful way of organizing many decisions crucial to the final quality of knowledge acquisition results and the effective transfer of the knowledge obtained. These decisions are made throughout the course of the knowledge acquisition effort, since many are contingent on new data that emerges as a result of the knowledge acquisition activity itself.

Major planning decisions in Canvas include the following:

- selecting the knowledge acquisition staff (investigators) and knowledge sources (informants),
- determining the roles of these participants, as audience, information sources, customers, etc.,
- anticipating, tracking and managing sources of bias in data gathering,
- managing variability in information resulting from differences of opinion, differences of viewpoint, or differences in work practice,
- managing access to source information, including problems of scheduling and confidentiality,
- managing access to new information, produced as part of the knowledge acquisition process, and
- selecting notations for representing the acquired knowledge.

Planning does not include the actual execution of knowledge acquisition activities such as reading documents, interviewing or observing (though results of any given activity may result in new information which requires adaptation of the plan). Thus, this guidebook does not provide detailed guidance in how to perform knowledge acquisition activities such as interviewing practitioners or examining artifacts. However, performance in the context of a systematic plan helps to guide certain strategic aspects of these activities (e.g., choice of topic) not addressed by specific techniques. This guidebook also does not provide detailed comparison of Canvas to other approaches to data gathering or knowledge acquisition (although it may aid the reader in drawing such comparisons). These subjects are potential topics for future papers and reports to complement this document.

The Canvas approach is directed mostly towards the problems of planning and managing a large-scale knowledge acquisition effort to support system engineering, expert system development or domain engineering (engineering of components for reuse in multiple systems). Canvas pays particular attention to the problems of transferring technical information, or transferring work practice knowledge from non-technical communities to technologists. Many issues raised in the document may be less relevant to primarily social or cultural studies, where the information being transferred is usually less technical in nature. Nevertheless, many of the Canvas principles (e.g., strategies for handling bias) are applicable to any situation in which knowledge is being systematically elicited for transfer from one work practice setting to another.

Although Canvas was designed from lessons learned managing a large project, many of the principles are just as applicable for small projects. Canvas provides a particularly useful perspective in cases where two or more professional cultures are involved in differing capacities, and the interactions among them must be carefully tracked. Canvas is also appropriate when the professional communities involved have complex structures; that is, they include several distinct groups with different stakes in the knowledge acquisition effort.

We do not consider the current document to be a stand-alone method or process model. While the guidebook reflects real project experience (e.g., the SEP-based TCIMS program, ODM pilot project efforts), it does not represent a distinct process that has been followed multiple times in order to yield a repeatable process description. Instead, it offers a set of useful concepts and guidelines that can be applied to knowledge acquisition planning.

1.2 Intended Audience

This document talks about knowledge transfer as an interaction between people in distinct communities of practice, each with shared and to some extent implicit contextual knowledge, language, culture and strategic objectives. The document itself is a collaborative effort between contributors representing different technical communities. Although we have attempted to present the material with as few presuppositions as possible about readers' backgrounds, those readers with experience and/or interest in one or more of these communities will find the concepts most familiar and of clearest value.

Canvas emphasizes knowledge acquisition (KA) as an aid in understanding the distinct communities of practice in which new technologies are conceived and created, and the workplaces in which these technologies are adopted. Canvas provides challenges to some conventional assumptions about how the information about a workplace can be transmitted to the people designing systems for that workplace, and gives detailed advice about how a knowledge acquisition project should be organized, based upon these challenges. We have attempted to write this guidebook from the "ground up", so that anyone who has given some thought to technology adoption issues can follow the concepts described here. However, the audience most likely to have initial interest in these issues are people who have had responsibility for transferring new technology into a workplace, who would like to examine the dynamics of knowledge acquisition in more depth as one explanatory perspective for why such attempts succeed or fail.

More specifically, Canvas offers particular benefits to readers having one or more of the following organizational roles:

- *System/software project managers* — For project managers, Canvas shows the benefits of systematic knowledge acquisition, and outlines the risks incurred when these aspects are not adequately considered in overall project planning. It also provides detailed advice about how to plan knowledge acquisition activities and track knowledge acquisition workproducts in the context of a broader engineering project, and suggests criteria for selecting and training knowledge acquisition staff.

- *System/software developers* — Canvas illustrates how the cultural differences in workplaces have an impact on the information needed for effective requirements engineering. System developers will gain an appreciation of the potential value to be gained from knowledge acquisition as an integral managed part of technology development. They will better understand why KA is a distinct discipline and the importance of having the work performed by competent practitioners.

- *Experienced knowledge engineers* — Canvas puts well-known knowledge acquisition techniques (interviewing, verification, representation) into a general planning framework for application on large projects. Knowledge engineers will see how their efforts are related to other KA activities. Depending on their level of experience, they might find new insights into the many aspects of KA treated by Canvas, including bias, variability, the specific challenges posed by performing KA in technical settings, or acquiring knowledge about multiple systems.

- *New knowledge engineers* — Canvas provides a concise description of knowledge acquisition, including its underlying assumptions, goals and required support infrastructure. This information can be of great help in learning what is expected of a knowledge engineer.

- *Domain engineers* — Canvas is derived in part from the ODM domain engineering method and is thus directly applicable to the data acquisition needs of the domain engineering process, especially of the sort described by ODM.

- *Social scientists* — Those people experienced in knowledge acquisition from a social sciences, academic or humanist background will gain better understanding for the distinctive challenges and opportunities of performing KA in a technology-rich environment. This will become of increasing importance for KA efforts as technology becomes a more ubiquitous presence in most cultural settings. In addition, Canvas addresses some ways in which technology can be applied directly to the support of KA activities.
- *Knowledge sources/informants* — Those who will potentially be in the role of informants or knowledge sources for a KA effort will better understand the process in which they are taking part. This will enable them to take a more pro-active and collaborative role, perhaps including observation of instances of investigator bias and articulation of these concerns.
- *Support technology developers* — Developers planning to provide supporting tools for knowledge acquisition will find detailed recommendations for how a repository of information must be managed in a KA project, as well as requirements for managing the other resources in the project.

For all readers, Canvas offers an awareness of how culture plays a role in knowledge acquisition and what can be done to address cultural issues. Anyone interested in knowledge acquisition or in making use of codified knowledge can use parts of this document to understand how knowledge can be systematically acquired and organized and why this is of value. Ideally, this could encourage more effective knowledge acquisition in many aspects of the workplace and cultural life where formal KA efforts are not likely to be initiated.

1.3 Guidebook Organization

The body of the guidebook is organized into the following sections:

— *Section 1: Document Overview (this section)* — Defines the purpose, scope, intended audience, and organization of the document and offers guidance in how to read it.

— *Section 2: Canvas Overview* — Describes the background and motivation for the project that developed Canvas, the context in which it was developed, and an overview of the knowledge acquisition planning process which is the scope addressed by the method.

— *Section 3: Canvas Core Concepts* — Articulates the conceptual foundations of Canvas, defining the boundaries around what is considered knowledge acquisition (for the purposes of this guidebook), and defining the basic terms used and implications of these foundations.

— *Section 4: Planning the KA Enterprise* — Provides a set of basic, practical guidelines for using Canvas to plan and manage a knowledge acquisition effort.

— *Section 5: Representation of Knowledge* — Elaborates on the role of representation in knowledge acquisition, and outlines some criteria that can be used to help knowledge engineers develop their own repertoire of representations. Section 5.0 also provides a sample repertoire, assembled from our studies of large knowledge acquisition projects.

— *Section 6: Dossier Planning and Management* — Provides detailed advice about how to structure a repository (herein called a "dossier") of knowledge acquisition products, including "starter sets" to help initialize the dossier.

— *Section 7: Conclusions* — Reviews the key principles embodied in Canvas and proposes possibilities for further work.

The guidebook also includes the following supplementary material:

- *Appendix A: Canvas as an ODM Supporting Method* — Shows how Canvas fits into the ODM process framework and can be used as an ODM "supporting method."
- *Appendix B: Representing the Knowledge Acquisition Process* — Uses Canvas on itself; the knowledge acquisition process is shown, represented in notations commonly used in Canvas.
- *Appendix C: Canvas Lexicon* — Definitions of the terms that are fundamental to understanding Canvas.
- *Bibliography* — Bibliographic entries for documents referenced in the guidebook and other documents related to Canvas.

1.4 How to Read the Guidebook

Each segment of the audience has different needs and interests and will thus benefit most from different portions of the guidebook. All segments of the audience should read Sections 2 and 3 (consulting the Canvas lexicon in Appendix C as needed) to gain a basic understanding of the method.

Readers with a background or interest in domain modeling will find Appendix A of particular interest. Developers of technology to support knowledge acquisition efforts will find Section 6.0 and Appendix B to be of particular interest. Knowledge engineers are likely to find Section 5.0 to be most useful for their needs, while project planners will need to master the ideas in Section 4.0.

The name and number of the current section is included in the header on each page to help the reader navigate through the document.

Treatment of key Canvas terms: In general, when a key term is first introduced in the guidebook, it appears in a ***bold italic*** typeface and is defined at that point. Such terms are also sometimes shown in bold italic when first appearing within other sections of the document to reestablish their identity as lexicon terms. All such terms also appear, with definitions, in the Canvas lexicon in Appendix C so they can be easily looked up whenever they are encountered in the document. Appendix B shows how many of the core terms relate to one another, and is itself an interesting case study in knowledge representation.

2.0 Canvas Overview

This guidebook presents an initial framework for planning and managing a systematic knowledge acquisition effort. The framework is organized in terms of a central metaphor, as suggested by the name "Canvas". One of meaning of the English word "canvas" is: "a coarse cloth of open mesh weave on which embroidery or tapestry is done." A knowledge acquisition effort (as defined for the purposes of this guidebook) requires the coordination of a number of people performing different roles in the work of sharing, gathering and transferring knowledge. Many conventional ideas about large-scale project management depend on the notion of decomposing a large task into small, relatively independent pieces. We believe the "canvas" metaphor is a better starting point for organizing knowledge acquisition activities, which are difficult to specify in advance and quite different in nature from technical design tasks.

The name of the Canvas framework is thus intended to evoke an image of weaving strands together into a coherent tapestry of knowledge. The strands begin as independent threads of information, guided by planners using the Canvas approach to form a woven fabric that has more structure and meaning than the individual threads. We use this metaphor throughout the guidebook when we speak of the interactions between life cycles of the various elements that participate in the knowledge acquisition process.

We refer to Canvas in different contexts as an "approach" (where the emphasis is on the key concepts and concerns/issues raised) or as a "framework" (where the emphasis is on the core set of terms and elements and their semantic relations). The framework is presented as a core set of concepts, planning guidelines, example representations, and a set of guidelines for initializing a ***dossier*** — a repository containing knowledge acquisition workproducts organized to facilitate access by the target audience.

This section describes the background and motivation for the project that developed the Canvas framework, and the context in which it was developed:

- Section 2.1 presents a motivation for the focus on knowledge acquisition in general, and on planning aspects specifically, from the standpoint of those involved with software technology development;
- Section 2.3 presents the specific goals and approach taken in developing Canvas;
- Section 2.2 presents the project context in which the Canvas approach was developed and documented;
- Section 2.4 describes the background of the case study which formed the basis for many of the examples in this guidebook; and
- Section 2.5 briefly describes some additional validation and experience.

2.1 Motivation

People acquire knowledge informally on a regular basis, as a natural outgrowth of performing work tasks, reflecting on past experience, communicating experiences to others and learning from others' experiences to improve practice. In many situations, however, it becomes important to follow a more systematic knowledge acquisition process. Familiar examples are rules for acquiring evidence in trial law or procedures used by journalists in gathering background information.

The notion of knowledge acquisition as a distinct phase of a technology development effort (and the use of the term "knowledge acquisition" to describe this phase) first gained prominence in the expert systems/artificial intelligence (AI) field, where the goal has generally been to codify expert knowledge in some domain into representations that can serve as a basis for automated deduction and decision support. Knowledge engineers in this field borrowed some techniques from other disciplines such as psychology, linguistics and the social sciences.

Within the broader software technology field, knowledge acquisition has not been as widely recognized as an element of the system development life cycle. Many similar activities are performed (e.g., interviewing practitioners and end-users, gathering information about the operational environment for the system to be developed) under the broad umbrella of "requirements analysis". Some borrowing from the same disciplines that contributed to the knowledge engineer's base of techniques has occurred.

It may seem paradoxical to speak of knowledge acquisition when there is no presumed goal of implementing an expert system. Yet experience has shown that in many situations the knowledge acquisition and codification process has yielded direct value to practitioners, independent of the eventual success or failure of the expert system initially envisioned ([6], [10], [58] and [44]). Similarly, because conventional software systems are often assumed to be automating routine activities, we may underestimate how closely entwined these activities are with practitioners' knowledge.

A unified approach to knowledge acquisition would be of benefit to a wide spectrum of technology development and adoption situations. Current techniques, however, are far from an adequate basis for such a unified approach. Some key concepts are needed, reflected in the framework presented in this guidebook:

- Knowledge is created and maintained in practice, and is often tacitly *embedded* in social interactions within the work setting;
- The activity of gathering, reflecting on, and codifying such knowledge, for whatever reason, is an *intervention* in the organizational dynamics of the work setting; and
- Knowledge acquisition involves the interactions of distinct communities.

The points above would apply very generally, even to knowledge acquisition carried out in an academic research context. There are significant implications, however, for technology developers.[1] If knowledge is embedded in practice, then any automation of people's work processes will intersect with significant knowledge-intensive activities. If requirements-gathering involves capturing aspects of that knowledge, then the *very act of gathering requirements* needs to be understood as an intervention in the organization. Last but not least, successful system development depends on knowledge acquisition processes as a kind of "culture contact" to effect transfer of knowledge between users and developers as distinct communities. Unfortunately, in too many instances this transfer fails, leading to significant breakdowns in communication and process, and inscrutable systems that solve the wrong problems.

Significant extensions to current approaches to knowledge acquisition are needed to address these issues. For example, in large software projects in which systems must be built to interact with

[1] This document reflects a focus on technology development, in light of the intended primary audience. Our main point of reference, and basis for experience, is *software system* development, which by its nature involves particularly close interaction with work practices where knowledge is embedded in the way outlined here. While the points apply more generally, we are attempting to provide as concrete a context for discussion as possible.

many legacy systems and work practices, comprehensive information about these interactions is imperative for proper integration of the software. But few approaches to knowledge acquisition are oriented specifically towards capturing the knowledge of users in technology-intensive environments.

One area of the software engineering field that has recognized the role of knowledge acquisition is in developing software for reuse. In this field, ***domain engineering*** refers to the analysis of a set or class of systems in a given domain, in order to guide the design of reusable software for that domain. Domain engineering presents two additional challenges for knowledge acquisition. First, the need to study *multiple* systems and environments of use means that variability information must be carefully managed. Second, in some respects domain engineering involves building software components to be used by software developers themselves, and therefore requires the capture of their expertise. Therefore, in a domain engineering project, software technologists themselves form one important practitioner setting that must be studied. Current approaches to knowledge acquisition are not well suited to gathering knowledge from these types of practitioners.

Systematic Knowledge Acquisition

The Canvas framework formalizes and clarifies a concept for systematic knowledge acquisition broad enough to address some of the issues raised above. ***Systematic knowledge acquisition*** involves repeatable procedures for making key decisions in planning and performing knowledge acquisition, and for recording results of KA activities in a way that preserves essential contextual information about the data acquired. A systematic approach to knowledge acquisition addresses the following issues (among others):

- *Sources of the information*: Where did the information come from and how was it obtained? Was there possible bias, misinterpretation or selective filtering on the part of the person who provided, or who obtained, the information?
- *Handling multiple information sources*: What is the relative convergence or divergence of opinion among the people providing information? Does this represent variance in opinion and belief or legitimate variation in the phenomena described?
- *Managing the data acquisition process*: This includes issues such as budgeting resources for data acquisition. Given scarce resources, unpredictability in accessing experts or the effort required for specific sessions, what are the most important information sources to consult? How do we know when we have acquired enough data? Handling of potentially sensitive or proprietary data must also be considered.
- *Access to the information.* Who will be using the information gathered, and how? Are there many audiences with differing perceptions about what the knowledge should be and how it should be used? How can we help all of these audiences find the information they need?

By focusing on knowledge acquisition *planning* we have excluded much specific detail about how to carry out knowledge acquisition activities such as interviewing or analysis. These are beyond the scope of the current document, but might well be expected as part of a comprehensive method for knowledge acquisition. The following section provides information on the project background which clarifies the rationale for the specific focus chosen for this guidebook.

2.2 Project Background

This guidebook is the result of a Defense Advanced Research Projects Agency (DARPA)-funded task involving collaboration between the Software Technology for Adaptable Reliable Systems (STARS) and the Health-care Information Infrastructure Program (HIIP) Programs. STARS participating organizations included Lockheed Martin Tactical Defense Systems, Organon Motives, Inc., and WPL Laboratories, Inc. HIIP participating organizations included ScenPro, Inc. and the University of Texas/Arlington. The overall objective of the joint task was to develop a software engineering method which integrates key elements of the STARS ODM and HIIP SEP methods. These methods are described briefly below, followed by additional rationale and motivation for the integration activity.

2.2.1 ODM Background

Organization Domain Modeling (ODM) is a method for ***domain engineering***, a discipline which extends conventional software engineering approaches by focusing explicitly on analysis across multiple application contexts to support systematic reuse. ODM is a highly tailorable and configurable domain engineering method, useful for diverse organizations and domains, and amenable to integration with a variety of software engineering processes and implementation technologies. The method offers a systematic, exemplar-based approach to analysis of commonality and variability, addressing analysis of both legacy systems and requirements for new systems to derive reusable assets focused within a domain. ODM grounds the domain modeling in the context of the organization and the relevant stakeholders.

Key features of ODM include the following:

- In addition to focusing on the strictly technical aspects of domain engineering, ODM emphasizes analysis of the diverse stakeholders that form the organization context within which each domain engineering effort is conducted.
- ODM provides systematic techniques for identifying and selecting highly focused domains of strategic interest within larger business areas, and for incremental and iterative scoping to mitigate risk and produce robust, coherent domain models.
- The ODM modeling life cycle details the transformation from descriptive modeling of legacy systems, artifacts and experience to prescriptive specification of architecturally integrated assets, designed for a well-scoped range of variability and characterized in terms of features relevant to domain practitioners.

  The resulting domain model offers a sound basis for making the design decisions and tradeoffs required to engineer reusable components that are robust, of high quality and intuitive.

ODM domain engineers study a carefully selected set of representative software systems within a domain. ODM provides criteria for selecting from a wide variety of knowledge elicitation techniques, including artifact analysis, interviewing of domain informants, and process observation. This might include analysis of artifacts from across the software life cycle, as well as use of ethnographic techniques such as those described in [42] and [43] (e.g., to capture process knowledge and undocumented "techlore" of application developers, users and other domain stakeholders). The resulting data is formalized in a domain model which represents common and variant features of systems in the domain. The process requires a unique mix of technical, conceptual and organizational skills on the part of the domain modeler.

Under funding by the DARPA STARS Program, ODM has been extensively documented in a guidebook [49] as well as in shorter papers [38] [39]. The guidebook provides explanations of key concepts, a formal process model (documented in $IDEF_0$ process modeling notation [21] [24] [41]), work product descriptions and templates, and detailed, practical guidelines for domain engineering projects.

2.2.2 SEP Background

The Scenario-based Engineering Process (SEP) [15] is a user-focused methodology for system development. The methodology has been applied and is currently in use in various health care programs, including the DARPA Trauma Care Information Management System (TCIMS) program. SEP has strengths in engaging the user in all phases of a project, and benefits from a well-structured approach to eliciting information. Utilizing scenarios as a means of engaging the user, SEP defines a series of representations that incrementally increase the formality of the information documented in order to provide effective transfer to technology developers. In terms of the background presented in Section 2.1, SEP represents a major step in integrating a systematic approach to knowledge acquisition into a technology life cycle which can span both conventional and intelligent systems development.

As shown in Exhibit 1, SEP consists of three processes: knowledge acquisition (KA), knowledge engineering (KE), and system engineering (SE), which form a well-defined semi-ordered set of procedures. The result of a SEP process is the construction of a component-based architected system.

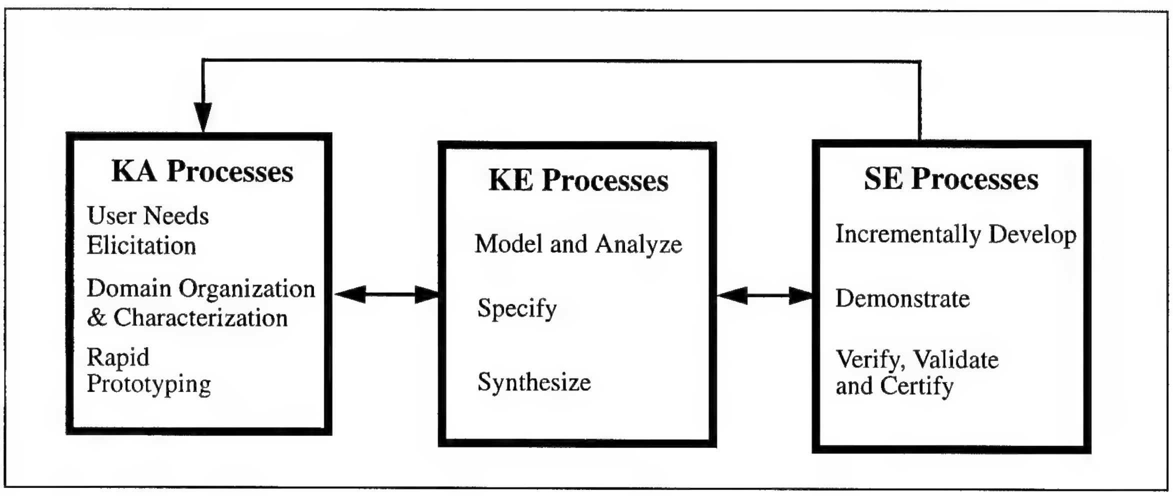


Exhibit 1. Scenario-based Engineering Process Overview

SEP has a number of objectives that distinguish it from other system development methodologies. These objectives are to improve communications among project stakeholders, facilitate the assignment of responsibilities, and maintain traceability within interdisciplinary system-in-the-large development efforts.

In order to achieve these objectives, SEP focuses on scenarios for knowledge acquisition and validation, iterative "build-a-little, test-a-little" prototyping, and a strategically planned incremental development approach. For the purposes of knowledge acquisition, and hence for Canvas, SEP's focus on scenarios is of the greatest relevance.

In SEP, a scenario is a single path through some work process, similar to case studies familiar from medical practice. In fact, in the TCIMS project, where SEP was applied to a medical domain, many of the scenarios read very much like case studies. Scenarios are the key components driving communication and traceability in SEP. Scenarios participate in the SEP lifecycle at several points. During knowledge acquisition, they are used to communicate with the prospective users of the technology, to determine their requirements and derive system specifications. Filtered and generalized scenarios result in task analyses, which, along with other knowledge acquisition products, are used during knowledge engineering for object-oriented design and architecture synthesis. Finally, the scenarios themselves are used again during testing and validation stages of system engineering to evaluate the developed system.

The ramifications of using scenarios as the basis of the engineering process are subtle but important. Scenarios are directly understandable to the prospective users, since they are simply records of their experiences; on the other hand, the same scenarios can be used by system developers as an "acid test" for software products, including specifications, requirements, and executing code. Using the same scenarios throughout the process facilitates traceability of workproducts. Since the scenarios are relevant to both the users and the developers (although in different capacities), they foster communication between the two groups. As opposed to other descriptions of the workplace (e.g., object descriptions, task structures, organizational charts etc.), scenarios provide direct information about how the practitioners in some workplace interact with other practitioners or systems, thus facilitating the identification of which practitioners or subsystems are responsible for which actions.

This final point is a critical aspect of the underlying motivation for SEP. The SEP method fosters the recognition that interactions of prospective users with their environment are flexible and dynamic. Static generalizations that talk about these interactions are useful for specifying and building systems, but in the final analysis, the systems must be responsive to the detailed, dynamic nature of the interactions. By basing the process on scenarios, this accountability is retained, since scenarios record interactions themselves in their raw form, without generalization. This ability to attend to the interaction between the end users and their environment is a primary value that SEP adds over other system development methodologies.

2.2.3 Relation of SEP and ODM to Canvas

Both the SEP and ODM methods are rooted in conventional software engineering approaches. Each method has diverged from those conventional approaches in significant, although different, ways which have made contributions to their respective areas. The methods also share many of the same concerns and offer complementary perspectives on key software engineering problems. There are thus substantial opportunities for synergy among the two methods. Exhibit 2 provides a schematic illustration of the relation of each method to its primary context of application, and the relationship between the two methods, as described in the following paragraphs.

As the SEP method summary above implies, SEP focuses primarily on engineering individual systems. Although workproducts produced in developing each system may prove of value in developing subsequent systems, the method does not emphasize multi-system analysis with the explicit objective of developing components that can reused in multiple application contexts. In contrast, ODM focuses explicitly on analysis of multiple systems to support systematic reuse. While many domain engineering approaches emphasize exploitation of commonalities across systems, a key feature of the ODM domain engineering approach is its rigorous analysis of variability across systems to support systematic management of alternatives within a domain. This rigorous treatment of variability is a key differentiator of ODM, and a key factor that requires extension of the SEP approach.

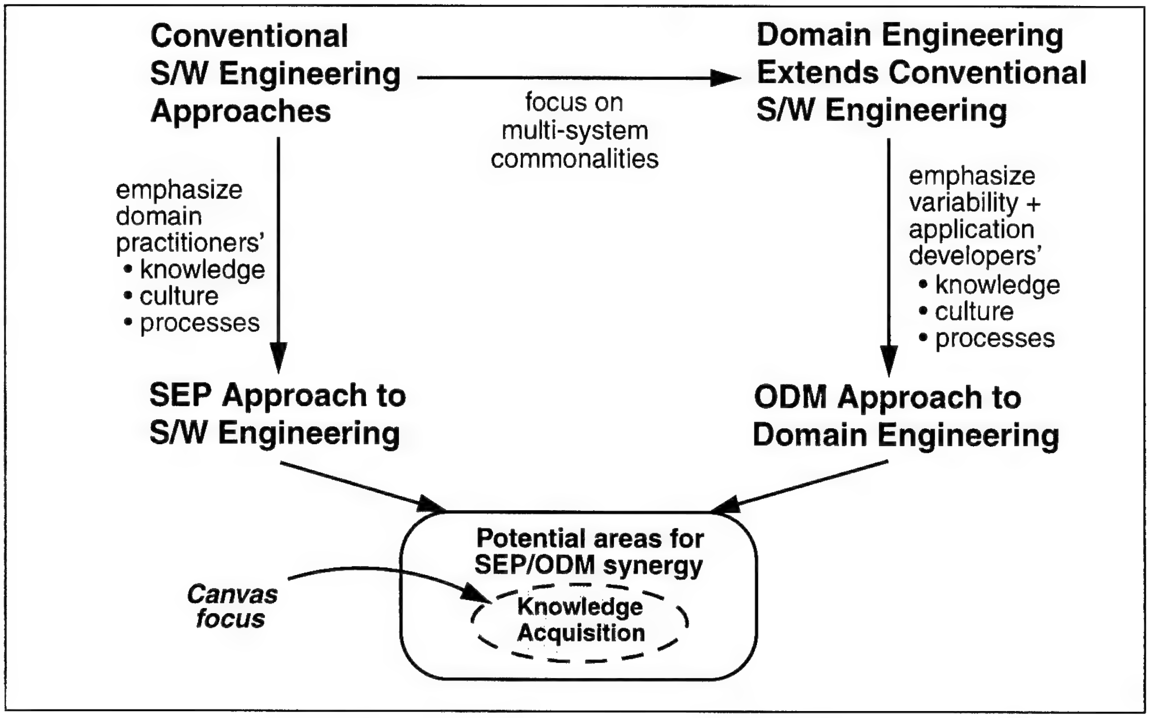


Exhibit 2. Motivation for a Composite ODM/SEP Method

In comparing SEP and ODM, it became clear that there were several areas in which the methods complemented one another and could benefit synergistically from cross-fertilization. These included modeling approaches and architectural concepts, among others. However, the area offering the richest integration opportunity was knowledge acquisition. Each method could benefit from a knowledge acquisition approach that unified the SEP and ODM perspectives (e.g., single system versus multi-system; domain practitioners versus application developers; scenarios/interviews versus artifact analysis.) It was from these insights that the Canvas approach was born.

Within their respective milieus (system and domain engineering), SEP and ODM both rely heavily on knowledge acquisition and employ a variety of knowledge acquisition techniques. The primary difference between the methods in this area is where knowledge acquisition techniques are applied, and the preferred methods. SEP focuses primarily on acquiring knowledge about the environment in which domain practitioners (i.e., end users) are practicing their craft. This is done mainly through scenario-based interviewing and related techniques. Knowledge acquired in this way helps system developers get an understanding of how work is done in that environment and gain insights into how it can be better automated.

ODM places stronger emphasis on acquiring knowledge about the developers' environment. (Domain engineering has been characterized as capturing and codifying the expertise of application developers in a given domain.) Because domain engineering is often performed in mature domains where there are legacy systems available for study, ODM also places more attention on the presence of technology in end-user environments. ODM also assumes that knowledge acquisition through interviews with practitioners (i.e., developers) will need to be augmented by analysis of artifacts describing existing systems (or requirements for new systems) in the domain.

2.3 Canvas Goals

Canvas has been specifically designed in response to knowledge acquisition needs perceived in ODM and SEP. A number of knowledge acquisition methods satisfy many of these needs. We have established the goals of Canvas to respond to those needs that, in our experience using these two methods, have not been fully met or adequately treated by other approaches:

- *Address the "cultural communication" aspects of knowledge acquisition explicitly in the planning process.* Both ODM and SEP pay particular attention to the interactions of prospective users within their environments, including interaction with automated systems and other practitioners. This places specific demands on the knowledge acquisition process, in that it must enable acquisition of knowledge about interactions among people with widely varying work backgrounds and/or systems designed by such people. One goal of Canvas is to make the study of interaction between different work settings central to the knowledge acquisition process.

- *Foster utilization of the full spectrum of available knowledge sources*, by providing a framework where different types of knowledge sources are considered as an integrated whole, not as a grab-bag of distinct types of data. Our experience shows that different projects tend to gravitate towards single modes of data collection, such as interviews, artifact analysis, or facilitated group sessions. An interview-oriented project, for example, will use artifacts or documentation more cautiously, as supplemental material or background briefings for interviewers. There may be a sense that interviews yield the "real story" whereas only the official story (policies, regulations, etc.) can be derived from documentary sources. Study of artifacts could be used more effectively in these settings via walk-throughs, validation of interview data, etc. In other contexts (e.g., some reuse efforts), there has been a tendency to under-utilize observation or interviews with people in developers' settings. Acquiring knowledge by cross-checking informant interviews, analyzing artifacts, and directly observing work practice provides a richer set of data and more possibilities for robust validation.

- *Emphasize legacy systems and anticipated new systems.* In many environments introduction of new systems or technologies must be preceded by some understanding of a large number of existing "stovepipe" legacy systems. (This is particularly true, for example, in the DoD arena.) Many of these stovepipe systems do not currently communicate or interoperate to an acceptable level. Knowledge acquisition to support technology development in these environments must have systematic ways of accounting both for anticipated new systems and existing legacy systems in the environments studied. In domain engineering, study of legacy systems is integral to the task, although the typical objectives are comparative analysis to support reengineering for reuse, rather than analysis to support introduction of new systems that can interoperate with, link, or replace legacy systems. From the KA standpoint many of the challenges are similar. A goal of Canvas is to extend traditional KA techniques to understand the role of legacy systems when addressing issues of technology development and use.

- *Validate/increase credibility of KA within technology community.* An important lesson learned in applying SEP is that system developers were not always convinced of the value of knowledge acquisition results, nor of the benefits of performing knowledge acquisition in a systematic way. We hope the Canvas approach will help demonstrate the importance of systematically planning and performing KA in order to meet system development objectives.

To meet these broad goals, the method integration effort addressed these specific objectives:

- Address acquisition of knowledge associated with legacy systems present in the practitioners' work setting;

- Address knowledge acquisition in technology development environments. Those who build systems have a unique picture of the work domain in which the systems will operate. This picture is distinct from the practitioners' view, but it has a powerful effect on the kinds of systems built; also, once those systems are fielded, the developers' picture will have powerful impact on the work practice itself.
- Address the systematic handling of variability information in the knowledge acquisition process.
- Generalize from the specific representation choices reflected in SEP (e.g., scenarios, task diagrams) and ODM (e.g., taxonomies for representing variability information) to a systematic approach to selecting and characterizing the appropriate representations to use in given situations.
- Incorporate stakeholder analysis into the KA planning process to help ensure that the overall knowledge transfer goals of the KA effort succeed.
- Investigate ways of rendering the knowledge acquisition planning and management process more efficient with automated support capabilities.

Applicability of the Canvas Framework

The Canvas knowledge acquisition approach is based not only on study and comparison of the SEP and ODM methods themselves, but on experiences gained by their application in a variety of contexts. In recognition of these lessons learned, Canvas is intended to be usable in a number of knowledge acquisition project contexts including but not limited to SEP and ODM efforts.

The first application is use of KA as an integral part of the software engineering life cycle. SEP experience on TCIMS and related health care applications suggest that KA techniques have an important contribution to make, even for systems with no significant AI component. KA techniques and the later knowledge modeling stages that translate KA results into more formal semantic representations provide a basis for more effective system requirements analysis. They can also be used to suggest new potential areas for automated support (i.e., pre-system concept development) and, at the back end of the process, can validate and evaluate use of a fielded system.

The second primary application of interest is use of KA techniques as part of domain engineering to support systematic software reuse. While the SEP method has contributed a repertoire of specific KA techniques and representations, ODM has provided a broader conceptual framework that encompasses a range of data sources, awareness of stakeholder issues in both the developer and user environments, and management of variability.

In extending a KA planning framework to cover these two application contexts, we have of necessity generalized it to accommodate other applications as well. Canvas principles (even though derived from SEP and ODM) are applicable in many other contexts; this document therefore does not assume that Canvas knowledge acquisition activities are taking place within a larger SEP- or ODM-based effort. (The discussions in this guidebook do, however, presume an overall perspective relevant to software system developers.) We hope that a wide range of projects can benefit from our integration of SEP and ODM experience and perspectives in the Canvas framework, by making more effective use of systematic knowledge acquisition as an integral aspect of the project.

2.4 Validation through TCIMS Experience

In developing the Canvas method, we have drawn extensively on the project experience of SEP method providers and other knowledge acquisition specialists in the Trauma Care Information Management System (TCIMS) project. The following paragraphs provide some context for this project, from which examples are cited throughout the guidebook, and explain how TCIMS experience was factored into the Canvas framework and this document.

TCIMS Project Background

The Trauma Care Information Management System (TCIMS) project is a two-year effort to demonstrate radical improvements to the nation's Trauma Care Information Infrastructure, both civilian and military. A consortium of 13 organizations under DARPA guidance is designing, and will demonstrate the benefits of, advanced computing and communications solutions to accomplish this goal. In pursuit of this goal, the TCIMS project has the following objectives:

- The completed system will be commercially self-supporting. Consortium members intend to develop and produce products that conform to the TCIMS reference architecture and are commercially viable in both military and civilian use, both urban and rural.
- TCIMS will improve field trauma care by providing relevant medical and patient information to field medical care providers, and transmitting patient data ahead to the hospital to minimize delays in scheduling emergency facilities and resources.
- The TCIMS consortium will develop national-level trauma care information management standards leading to rapid price reductions in the cost of such information and inter-operability among all users and providers.
- The consortium is developing a TCIMS architecture to promote continuing trauma care system development and innovation.

When medical care arrives at the scene of a medical trauma scene, a Personal Status Monitor (PSM) might be attached to the ill or injured person to sense their condition. In a military environment, each patient is expected to have an identification and information card such as the AT&T "Smart Card," which provides patient identification and a minimal medical history. The PSM will have a micro controller/memory module and a communication scheme. It will monitor the patient's vital signs and location, and will give an alert if the patient's condition becomes critical. The PSM is being developed by another DARPA program.

PSM and Smart Card data would be read by a Field Medic Associate (FMA) computer carried by the medic, which will provide a multimedia display of patient information. Voice command is among the interfaces being considered to input information to the computer. The Field Medic Associate will also provide the field medic with access to clinical knowledge, information, and advice needed to treat each patient. The FMA will allow the medic to gain on-line information about resource and facility availability. The individual Field Medic Associates will feed information into a Field Medic Coordinator computer, which will give a situation overview for crisis managers and aid in focusing attention on the most critical patients.

The Field Medic Coordinator computer will in turn feed information into the trauma center's Trauma Center Coordinator computer in the Clinical Workstation System, so that a complete and current case history on each patient will be built at the hospital even as the patient is being treated in the field. That history will follow the patient through the trauma care system. The Trauma Center Coordinator will provide rapid, complete information to the Patient Care Units and to portable

Trauma Care Associate computers carried by care providers in the hospital. The entire TCIMS will allow direct communication among medical decision-makers.

TCIMS will, therefore, provide rapid treatment procedures to help the medic make treatment decisions in complex and confused situations. The medical center will have access to on-line patient records, including x-rays and other test results. TCIMS will reduce time to treatment, will bring knowledge and information to decision-makers, will increase ratio of treatment to processing, and will reduce overall patient processing time. Finally, TCIMS will link clinical, tactical, and strategic planning and management.

Role of TCIMS data in preparing this guidebook

In the process of developing Canvas, the authors studied the TCIMS Knowledge Acquisition Plan for TCIMS in detail and debriefed Lisa Mantock, one of the project's lead knowledge acquisition specialists, on the planning process. We also observed a TCIMS-related interview session with a helicopter pilot involved in medical emergency transport, and worked with scenario data, domain models represented in the Loom knowledge representation system, and an on-line repository of TCIMS knowledge acquisition data (SEPWeb). Most of these materials are proprietary to the TCIMS consortium members and thus have not been cited in the Bibliography section or included directly in examples within the text. Nevertheless, the material provided one extensive data point for the framework and related recommendations offered in this guidebook.

The TCIMS project represented a large-scale knowledge acquisition effort within an even larger, complex consortium of technology developers and researchers. As such, it provided an extremely rich case study. It is worth noting a few aspects of the overall technology goals that made a knowledge acquisition approach particularly valuable. Although not an expert system project per se, a large motivation of the TCIMS effort was to introduce technology that would make detailed medical knowledge available via the system to assist practitioners and decision-makers in the field. In addition, the nature of the trauma care domain highlights the importance of communication and collaboration scenarios involving high-context knowledge on the part of the actors. The opportunity to incorporate lessons learned from the extensive KA work undertaken as part of the TCIMS project has enriched the Canvas framework and ensured that it is grounded in experience in (at least) health-care related domains.

2.5 Additional Validation

After completion of Version 1.0 of this document, some additional experience was gained working with the ideas in Canvas. Lockheed Martin and Organon Motives staff conducted a brief trial application of the Canvas planning process. In addition, Organon Motives created training materials based on key Canvas concepts and worked with a commercial customer developing software for the higher education administration market. We had the opportunity to work with customer staff who had been conducting a series of knowledge acquisition sessions as part of the overall product development effort. In addition, training was delivered to the "customer's customer", an in-house team at a higher education facility that intended to conduct some knowledge acquisition sessions as an internal effort. In the latter case, the trainees were not software engineers by background.

Unfortunately, the results of these applications of Canvas were not available for direct incorporation into this revision. Some individual insights are reflected in modifications or extensions, but are not specially indicated. In addition, this additional experience aided us in weaving many individual themes introduced in the earlier version into a more integrated approach. We hope to validate this material more extensively in the future and to continue to apply it.

In addition to these applications, Jim Solderitsch of WPL Laboratories used the archiving tool OpenRLF [50] to create an on-line index of publicly available TCIMS knowledge acquisition materials. The Canvas recommendations for building such an index (on which this effort was loosely based), as well as a more complete description of the on-line index, can be found in Section 6.0.

3.0 Canvas Core Concepts

This section introduces core concepts necessary to understand the Canvas framework for planning and managing the knowledge acquisition (KA process). Section 3.1 provides an explicit set of defining features for the general phenomenon we call knowledge acquisition. Section 3.2 introduces the key elements of the knowledge acquisition "canvas" as a central metaphor for the approach in this guidebook. Section 3.3 extends this basic framework to address issues central to addressing the Canvas goals of providing a systematic approach to knowledge acquisition that is to be integrated with system or domain engineering projects. These issues include knowledge acquisition in technology-intensive settings, management of variability, and dealing with the dynamics of KA as organizational intervention.

The core concepts introduced in this section will become the essential building blocks used in creating a systematic plan for the knowledge acquisition aspects of a project. We use the term ***knowledge acquisition enterprise*** to refer to a knowledge acquisition effort coordinated with a systematic plan of this kind. Section 4.0 lays out the structure of such a plan and recommended steps for its creation and ongoing maintenance.

3.1 What is Knowledge Acquisition?

We view ***knowledge acquisition*** as a special case of a much broader area of human activity we will refer to simply as ***knowledge creation***, activities that result in new knowledge being created. Individual learning, formal teaching, process capture, research, and knowledge acquisition or KA (the focus of this document) are all forms of knowledge creation. This document presents a particular approach to KA, the Canvas approach, which places emphasis on certain types of settings in which KA can be performed and a core set of concepts to consider in performing KA. The general relationships between these areas is depicted in Exhibit 3.

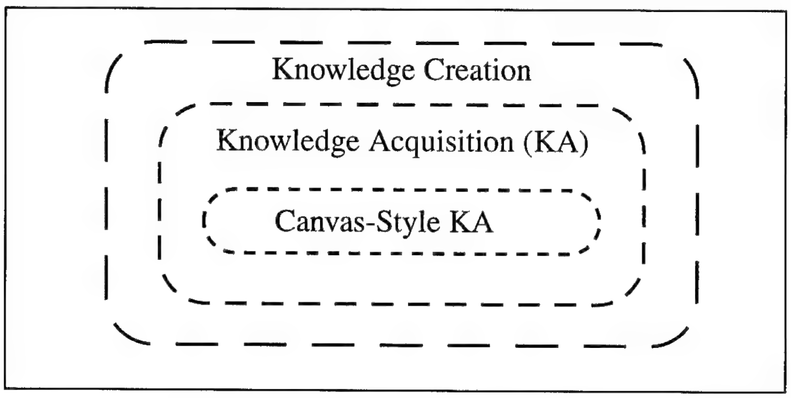


Exhibit 3. KA as a Special Type of Knowledge Creation

The following sub-sections elaborate our provisional definition of knowledge acquisition (KA):

- A central set of concepts necessary to differentiate KA from these other forms of knowledge creation involves the notion of a community of practice, outlined in Section 3.1.1;
- Section 3.1.2 builds on these concepts to outline the defining features of KA;
- Section 3.1.3 introduces the distinct community of practice roles that characterize the KA process;

- Section 3.1.4 provides an overview of the different forms of knowledge elicitation used within KA;
- Section 3.1.5 clarifies the distinctive aspects of KA by comparing it to other, more familiar forms of knowledge creation; and
- Section 3.1.6 introduces some high-level "modes" of KA in terms of the primary objectives of the activity and audience for the workproducts produced.

3.1.1 Communities of Practice

We will use the general term ***work setting*** (or simply ***setting***) to denote any environment where people interact with each other and perform work processes. Consider, for example, the setting of an emergency room at an urban hospital. Events occur in the work setting; some are anomalous and have little to do with the nature of the work being performed (e.g., a power outage); other events are routine or recurring processes essential to the overall mission of the work performed in the setting (e.g., taking a patient's information, checking insurance coverage, informing doctors of patient status, administering treatment).

A variety of people move into and out of the setting as participants or performers in events. Some of these people, such as the patients and family or friends that accompany them, may enter the setting for a specific event, and may know much or little about the work going on in the setting. But in almost all work settings (or, more descriptively, ***work practice settings***) a specific set of people, termed ***practitioners***, have ongoing roles within the setting. The notion of a work practice setting thus implies a certain stability, in that the same practitioners work together on a routine or regular basis, performing the essential processes of the setting.

For the purposes of our focus on knowledge acquisition, it is this work ***practice*** and not merely the work processes per se that are of interest. The reason is that knowledge is created through repeated performance of tasks, experience of classes of events, and through interactions with other people. Although a rigorous, formal definition of "knowledge" (as opposed to data or experience) is not necessary to establish these core concepts, we are interested in these aspects of knowledge:

- Much of our knowledge is created, and embedded, in *action*. Therefore, people cannot always articulate their knowledge verbally and out of the context of the work being performed. A nurse that has taken blood pressure hundreds of times may not accurately report what s/he actually does in the process; s/he may instead report the "official" routine that is to be followed. Some observation, or other techniques to place the knowledge in the context of action, may be necessary to effectively elicit accurate and complete knowledge.
- Much knowledge is *informal* or *tacit* in nature. "We know more than we know that we know." Tacit knowledge is what allows work practice settings to operate efficiently; people do not have to continually re-establish context among themselves as they perform their tasks. But this very tacitness which creates efficiency and "seamlessness" in the work process can create challenges in eliciting knowledge and transferring it to others that are not practitioners.
- Much knowledge is *socially constructed and maintained.* Since we collaborate in work practice settings, not all the knowledge important to effective performance is held by a single individual; some is held in the learned and practiced interactions between performers. This kind of knowledge may also be elicited (or at least identified) by observation, or by situations where people reflect in group settings on their practices.

We use the term ***community of practice*** to encompass both the community of practitioners (the individuals involved) with their work and associated social interactions, and the knowledge that is held by that community, individually and collaboratively, explicitly and tacitly, accessible in formal concepts and descriptions or embedded in action. (Often in this document, we will refer to communities of practice simply as communities, where the context makes the usage clear.)

Communities of practice are ways of grouping people that share language, experiences, common settings of action, social interaction, technical knowledge, and strategic interests. Members of a community of practice can effectively exchange information that depends on a high level of shared and assumed context. The concept implies that people from outside the community will have difficulty in utilizing the same information in the same way, participating in key activities that define membership, or in properly interpreting workproducts used in community settings.

Communities of practice are related to, yet independent of, specific work settings. A given work setting may reflect the interaction of multiple communities of practice (the doctors, transport personnel, administrative staff working together in the emergency room setting); and communities of practice may span numerous work settings (doctors have shared practice in emergency rooms, surgery, hospital staff meetings, etc.) The notion of a community of practice also has some obvious overlap with what might be termed general corporate or organizational culture. Groups of people working together in organizations do form cultural norms and customs. However, the emphasis here is on social interactions and broader cultural issues only insofar as they affect where and how knowledge is created, shared and transferred.

3.1.2 Defining Features of KA

Our definition of knowledge acquisition builds directly on the community of practice concept. We are particularly interested in knowledge ***embedded*** in a community of practice; that is, knowledge tightly interwoven with action, hidden behind tacit or unarticulated assumptions, or maintained through collaborative as well as individual practice. Although we deliberately leave the base term "knowledge" itself undefined in our lexicon, the approach here is roughly aligned with the concepts of knowledge and learning from the "learning organization" field; i.e., knowledge as the capacity for effective action in a given domain, where effectiveness is assessed by a community of practitioners (see [34], [20]).

Since this kind of knowledge is often transferred informally to new practitioners (e.g., through apprenticeship), problems arise in attempting to transfer such knowledge to new work settings. Many formal processes and techniques in knowledge acquisition are designed to deal with this specific set of challenges. The Canvas approach builds on and elaborates these approaches to KA, and extends them to accommodate special requirements of KA for system and domain engineering. In particular, a community-of-practice perspective can greatly enrich techniques for knowledge acquisition to support technology development and adoption goals. This is discussed further in Section 3.3.2.

Within a community of practice, we can distinguish three levels of "knowledge" work:

1) *routine* work activities performed by individuals and collaborating groups of practitioners;

2) knowledge *transfer* and propagation activities: training new people to be practitioners, exchanging knowledge among experienced practitioners, renewing and improving skills and practices, veteran practitioners codifying their knowledge;

3) activities that *create new knowledge*: observations leading to experiment, trials of new methods, formal research, etc.

Every community of practice has established procedures and conventions for each of these levels. In the medical field, routine work activities might involve many kinds of data-gathering: intake personnel gather initial patient data; nurses and doctors take more detailed patient histories; patients are informed (to varying degrees) about relevant conditions, given instructions on how and why to take medications or to follow procedures, etc. In terms of knowledge transfer, there are established procedures for education, apprenticeship, internship, coaching and supervision. Knowledge creation activities involve substantial ongoing research initiatives, publications, etc. Established "knowledge work" processes can exist, not only within, but also between and among communities of practice. For example, the medical field has a variety of ways of providing education to the general population. This represents knowledge transfer across distinct communities.

For the purposes of this document, we characterize ***knowledge acquisition*** (hereafter abbreviated as KA) as a knowledge creation process that makes a shift or *intervention* in the established systems for knowledge deployment, transfer, or creation within one or more communities of practice. This intervention can occur in one or both of these ways:

- A process that shifts knowledge from the status of tacit or embedded knowledge to that of ***codified*** knowledge, and/or
- A process that transfers knowledge to new communities of practice.

To qualify as KA, the activity must be recognized as distinct from the processes at any of the three levels of knowledge work outlined above, *from the perspective of members of the community of practice itself.* That is, we define "intervention" as socially constructed: i.e., if the community in question interprets the results as a change to the knowledge transfer systems in place prior to the KA initiative. In the first case, this can happen because the codification of the knowledge makes it amenable to new mechanisms for knowledge transfer. In the second case, transfer to a new community of practice (as opposed to an individual) means that the knowledge has been re-cast in a way that opens up new possibilities for knowledge transfer.

Thus, KA represents a kind of "second-order" learning on the part of one or more communities of practice. The focus of interest is on knowledge created and maintained within a community; the goal is to codify such knowledge in a new way that facilitates its transfer to multiple members of a distinct target community, on an ongoing basis.

3.1.3 Community of Practice Roles in KA

Our definition suggests that there may be one or many communities of practice involved in a KA effort. It will be helpful, however, to distinguish three distinct *roles* played by communities of practice: focus, investigator, and target communities:

- The ***focus community*** is the community of practice which holds some knowledge that is the ***focus of interest*** for the KA effort.
- This knowledge is of interest to some stakeholders in the ***target community***.
- An individual or team from the ***investigator community*** is directly responsible for carrying out the KA activities to effect the knowledge transfer from focus to target communities.

A single community may play one or more of these roles, and multiple communities may play a similar role. In particular the target community might be the focus or the investigator community, or perhaps a third distinct community. Relations between these communities are discussed in more detail in Section 3.1.6. The important concept here are the distinct *roles* played by each community.

Three distinct elements qualify a process as knowledge acquisition; the knowledge must be ***elicited***, ***codified*** and ***transferred***.

- Knowledge is ***elicited*** from a ***knowledge source*** drawn from the focus community. This elicitation process involves the active participation of a practitioner from the investigator community, an ***investigator***. After elicitation, some knowledge is available to the investigator that was not available before elicitation. The elicitation process involves interaction between the focus and investigator communities.

- In the next phase, some of this elicited knowledge is ***codified*** in a ***knowledge acquisition workproduct***. Codification may involve a number of processes, such as refining the documented knowledge from informal to formal representations, or synthesizing and summarizing knowledge from a variety of sources. This codification is a primary responsibility of the investigator community. It is important to note, however, that distinguishing elicitation and codification as separate processes does not rule out their taking place in a unified way. Elicitation and codification may co-occur as a ***collaboration*** between knowledge source and investigator; in fact, many aspects of the Canvas approach are oriented towards encouraging and facilitating this collaborative form of knowledge acquisition.

- In the final phase of the KA cycle, codified knowledge is ***transferred*** to some members of the target community. This phase involves active participation of members of the target community. It may involve direct contact between investigators and the target community, but the transfer is more than just individual to individual. It is the transfer of *codified* knowledge that completes the KA cycle, not merely transfer of knowledge that may have been acquired by the investigator during the elicitation process as individual learning.

3.1.4 Forms of Knowledge Elicitation

Given the elements introduced so far, we can usefully distinguish several forms of knowledge elicitation, each introducing different variables and issues that need to be considered in the planning process. As mentioned above, knowledge elicitation involves interaction between investigator and knowledge source. A ***knowledge acquisition session*** is an elicitation activity that involves at least one investigator and at least one knowledge source.

- A person serving the role of knowledge source in a KA session is termed an ***informant***. Typically an informant is a practitioner in the focus community. Knowledge can be elicited from informants in a variety of ways, including one-on-one face-to-face interviews, facilitated group sessions, walk-throughs and demonstrations within the workplace, solicited input via surveys, etc.

- A document or other workproduct, created within and for use in a work setting of the focus community, but serving as the knowledge source in a KA session, is termed an ***artifact***. Since we are particularly interested in knowledge that is tacitly embedded in the focus community, the artifact should play a distinct role in the focus community. For example, investigators might read a user manual, a field report, an article in a newspaper, etc.

Investigators can also directly observe or participate in work processes in a setting of the focus community. Observation is particularly useful as a cross-check of the other two forms of elicitation, in settings where there are few artifacts to study, or where informants cannot easily articulate the nature of their work. Also, one brief observation very early in the KA planning process can be extremely helpful in planning what artifacts to study, who to interview and how to interview them.The fundamental types of sessions are thus ***studies*** of artifacts, ***interactions*** with informants, and ***observations*** of events or work processes in focus community settings. (These are provisional

distinctions; there are clearly ways of gathering information that constitute borderline cases, e.g., observing a practitioner interacting with a system.)

Because we have defined KA as an intervention in the focus community's established patterns of knowledge work, a KA session is distinct from a normal work process taking place in a focus community setting. When a person acts as an informant, they are performing a different role, in a different setting, than their customary role as practitioner. Similarly, use of a workproduct from the focus community as an artifact is different from the original purpose for which the workproduct was created; it is being studied in a different context, the knowledge acquisition context. When an investigator observes events in a work setting, the setting doubles as a work practice setting and the setting for what is effectively a KA session. This usually changes the dynamics of what happens in the setting.

To qualify as knowledge elicitation, a session must result in the investigator gaining some knowledge about topics of interest and codifying this knowledge in some form; thus some knowledge acquisition workproduct is created as a result of (possibly during) each session. Every KA workproduct created in this way has an intended ***audience*** for whom it is intended within some specific community of practice. The workproduct may be further refined in later sessions, in this case the workproduct may indirectly play the role of the knowledge source for those sessions.

Most KA workproducts that are in final codified form will have the target community as audience. Knowledge has been effectively codified with respect to the target community when it has been documented in a workproduct with that community as audience, and if members of its intended audience can acquire the knowledge by examining (reading, viewing, running, etc.) the workproduct. KA workproducts can also be created with an intended audience of the focus community (e.g., for validation) or the investigator community (e.g., coordinated knowledge sharing within the KA investigator team).

3.1.5 Distinctive Aspects of KA

Having presented the defining elements of the KA process, it may be helpful to examine how KA activities would be distinguished from a number of the other modes of learning or knowledge work discussed above.

It is helpful to keep these distinctions in mind for several reasons. It helps clarify the intended scope of the planning approach provided here. It also identifies a number of associated processes that do take place along with the primary KA processes of elicitation, codification and transfer, hence must be anticipated and handled by the planning and management process. In addition, since KA activities are not part of the normal course of events within a work setting, practitioners will attempt to understand the KA process in terms with which they are familiar. For example, a KA interview might be interpreted by an informant as "teaching the interviewer how to do it". Understanding the relationships, and distinctions, between KA and these related forms of knowledge creation will help, therefore, in determining how to best present and characterize the KA effort within the focus community.

To provide context for the discussion, we return to the example hospital emergency room setting.

- *Information Transfer.* Within a work practice setting, people routinely exchange information in order to get their jobs done. For example, hospital administrative staff request information from incoming patients. Here, information is transferred between participants in the work setting to facilitate the routine work activities that are the focus of that setting. Such information transfer fails to qualify as KA because it is confined within a single work setting, and because the transfer is considered part of the routine practice in that setting.

- *Experience.* As people perform repeated tasks in a given work setting, they gain experience with that work. Exposure to repeated situations and variations in the conditions encountered increases their competence over time. For example, a nurse may learn from experience a variety of techniques for administering shots to patients who exhibit fear of injections. Though they are "acquiring knowledge" through experience this kind of learning is also, strictly speaking, out of the scope of knowledge acquisition as we define it here. Such experience has not yet been rendered transferable to others through codification.

- *Reflection.* As people gain experience, they often actively reflect on that experience and create new knowledge through this process of reflection. The reflective person gains knowledge through an activity that is often not part of the normal work practice, involving remembering and reviewing activities outside of the work context, or just a brief period of "staring out the window" time. For example, a nurse may reflect on techniques for calming patients by comparing what seems to have worked for children vs. adults, men vs. women, etc. Reflection does not yet qualify as KA, since the knowledge has not been transferred, nor even, more importantly, *rendered more transferable*. The new knowledge still affects the individual, not their overall community of practice. No attempt has been made to cross from one community of practice to another.

- *Teaching/Learning.* The processes described above—work practice, repeated experience, reflection, codification—are all aspects of individual learning. When new people are brought into a work practice setting, they generally go through an extensive learning process to become effective practitioners. Knowledge can thus be effectively transferred from individual to individual. For example, suppose our nurse is given the job of "breaking in" newly hired nurse trainees, and begins routinely coaching them on the fine points of calming down fearful patients before injections.

  This is a classic example of how knowledge is transferred informally within a community of practice. It does not yet constitute KA; in this case, the individual nurse must transmit the knowledge directly. Transfer of knowledge from an individual, as an individual, is a form of *teaching* but not knowledge acquisition per se. In a complementary sense, knowledge transferred to an individual could be considered *learning* on the part of that individual but not knowledge acquisition per se.

- *Writing Down.* As reflection and experience become more formal people may write down or codify their knowledge. The act of "writing it down" is an essential element of knowledge creation. Codification could involve something as simple as a checklist of items to remember each time a process is performed, something more formal such as a procedures manual, or the automation of the process through a software application. For example, suppose our thoughtful nurse creates a "Helpful Hints for Patients Who Hate Shots" checklist.

  The more a work environment encourages reflection, codification and sharing of knowledge among co-workers (i.e., the more the environment takes on the culture of a learning organization ([34], [31]), the easier it will be to shift tacit and informal knowledge to codified knowledge that is more easily transferred. Nevertheless, whether the write-up is intended for personal use, or for the use of professional colleagues within the same community of practice, the activity does not yet involve KA, since no effort has been made to *transfer* from one community to another. Since readers are being brought into the same community of practice as the codifiers, no cross-community transfer occurs. (Explicit attempts to write up material for other work settings may begin to take on some aspects of knowledge acquisition.)

- *Transfer within a Community of Practice.* Word gets around about our nurse's highly effective "bedside manner" and requests for training from other parts of the hospital staff are received. The checklist becomes a more formal training module of an orientation course for new nursing staff. From the standpoint of our framework, this does not yet constitute KA

because the forms of knowledge transfer are still those pre-existing in the community of practice; the target community is still that of practitioners (although novices) within the work setting. The knowledge has not been made available to a *new* community of practice.

- *Research.* A staff psychologist gets interested in the nurse's techniques, and obtains a grant to run a series of experiments to test a variety of techniques on different types of patients. The result of this research is new knowledge that was not previously available within the community of practice. Curiously, this activity would not fit within our framework of KA because it has the express purpose of creating new knowledge rather than eliciting knowledge that is already present (if tacit, or embedded) in the community of practice.

- *Transfer to an individual in another community of practice.* If an interested party from some other community of practice reads a book that is accepted as an authoritative source, and learns these ideas, then knowledge has been transferred from one community to another. This activity is commonly mistaken for KA; however, the ideas have not been made more accessible to the new community in general, only to a single member of that community. A similar situation holds if the interested party talks to the expert or attends a lecture, and thereby learns the ideas. Personal transfer can be made by pursuing a personal change to a single member of the community; transfer to an entire community involves finding a way to present the knowledge in a way that is accessible to members of the new community in general. If the interested party accomplishes this, then knowledge acquisition has indeed occurred.

Each of these learning or knowledge creation activities is certainly valid and useful in its own right. In saying the activities do not fit within the Canvas framework for KA we imply only that this framework is not targeted towards management of these types of activities: i.e., they may not require all elements of the Canvas planning approach, and may require other elements not addressed here. Many of the core concepts presented here might still be useful in improving these knowledge creation activities. (For example, the nurse might benefit from knowledge of KA techniques in first interviewing patients about their past experience with injections.) But these are incidental to the main focus of the approach.

Let us consider a few possible scenarios, given the example traced above, that would constitute a knowledge acquisition situation in the Canvas sense:

1) At the urging of the staff psychologist, a staff specialist in the new field of "Medical Ethnography" has just been hired at the hospital. Hearing about our now-famous nurse, the specialist gets a grant funded to do a comparative study of the way that a representative sampling of experienced nurses at different hospitals administer injections. The research includes in-depth interviews where nurses are asked to describe their procedures, the rationale for their use of different techniques, and their ways of classifying patients with respect to predicted reactions to the injection procedure. The research also includes examination of written procedures and regulations, some direct observation and videotaping of nurses giving injections, and comparison of reported to observed procedures.

2) In a follow-on study, a sampling of patients are interviewed about their past experiences with nursing staff administering injections. Some interviews are coordinated with observations from the previous study. The study focuses on patients who have required frequent administered injections from many health care providers. For various reasons, children below the age of twelve are not interviewed directly; their parents are asked questions about reactions they have observed.

3) A medical technology company is developing a machine that will be able to automate certain aspects of injections in the hospital setting. Results of the former studies are made available; but it is determined that additional information would be necessary to appropriately identify

requirements and potential risks of the envisioned technology. This leads to the initiation of a more extensive information-gathering effort which includes interviewing nurses about their experiences with other technology that has replaced manual procedures, their concerns and their suggestions for the proposed technology.

What are the aspects of these scenarios that constitute KA in terms of the definitions previously provided? There are a number of points worth noting:

- Scenario 1 is oriented toward collecting knowledge from a distinct group of practitioners. The nurses interviewed are informants because their information reflects more than their individual knowledge. They are interviewed as *representatives* of their community of practice. If the focus of a knowledge transfer process is a single individual's specialized knowledge, that process need not involve issues of the individual considered as an *informant* with respect to a given community of practice. An informant has a unique history and range of experience with respect to the settings within which the knowledge of interest is created, and participates in a web of relationships with other members of the community. Furthermore, the KA process intervenes in the community (at least indirectly) through the sessions involving the informant.

- Furthermore, the objective is to capture and organize the nurses' *existing* knowledge. If they are interviewed in groups, or get to see results of other interviews, or reflect on their own practice in new ways as a result of the interviews, they may obtain new insights about their own knowledge in this topic area. But the design of the project is not oriented toward this as the primary goal.

- Scenario 2 may at first appear to be about gathering "experience reports" rather than knowledge per se. However, since knowledge is created from repeated and comparative experience, patients who have received many injections have a unique perspective to offer on the same activity. They may classify health care providers according to differences in their "bedside manner" (e.g., brusque, informative, sympathetic) just as the nurses classified patients' reactions. This is a dramatic example of why we prefer the term "informant" to the term "expert" or "domain expert" prevalent in knowledge acquisition performed within an expert systems development context. Not all informants need be those labelled as "experts" within a given community of practice. In fact, informants need not even be the practitioners in the work setting, if other participants have a basis for knowledge that could be acquired and usefully deployed.

- Scenario 3 points out the importance of the topic of interest as a key element of a knowledge acquisition effort. The earlier studies had the objective of improving work practice in the hospital environment by disseminating "best practice" in a more systematic way. The technology development effort had different objectives, hence needed to elicit knowledge from the same community of practice (nurses) about different topic areas (experience with other technology in the workplace).

A few other points are worth noting. The scenarios demonstrate the role of informants, artifacts, and observation of work practice, and strategies for using these various modes of knowledge elicitation to cross-check and validate each other. The scenarios also demonstrate that the investigators play the role of a distinct community in the KA effort. Unlike the informants, their role is *not* predicated on their detailed knowledge of the topic of interest; however, they are expected to know something about the knowledge acquisition process (e.g., research protocols).

Finally, the scenarios demonstrate the two types of interventions that we claim characterize a KA effort. Scenarios 1 and 2 could both be considered interventions in the mechanisms of knowledge transfer within a given community of practice (the hospital environment); Scenario 3 represents the need to make a body of knowledge accessible and transferable to a different community of

practice. The objective in this case is not to train software developers how to administer injections, but to aid them in designing a system that will automate certain aspects of that process in a way useful to practitioners. Section 3.3.2 will further explore the idea that software technology always represents, to some extent, a way of codifying practice within a given work setting; hence, that a KA process is inherent, if often implicit, in the software development life cycle.

3.1.6 Knowledge Transfer Modes

Given the definition of KA above, we can identify several distinct ***knowledge transfer modes*** or patterns of knowledge acquisition. Each transfer mode corresponds to one of the three basic configurations of the knowledge acquisition process shown in Exhibit 4. In all three configurations, the elicitation process involves the focus community as a knowledge source and depends on codification performed by the investigator community. Each configuration can be varied by introducing multiple instances of any or all the community roles. The configurations differ based on the intended target community:

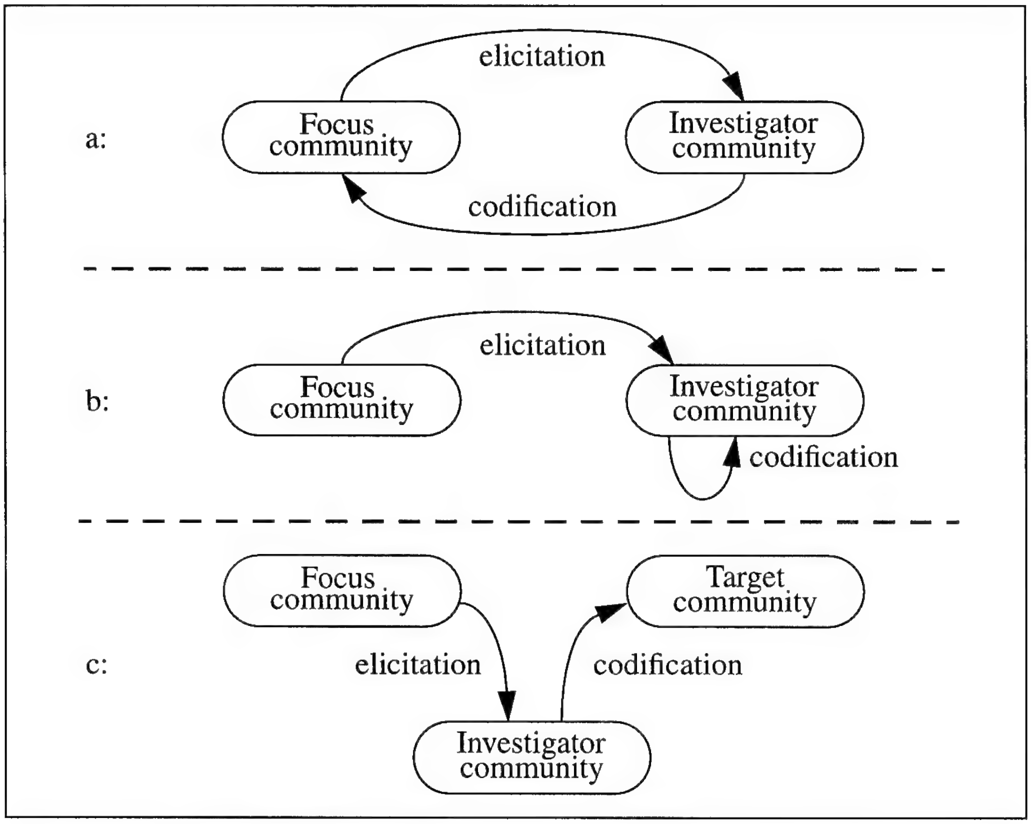


Exhibit 4. Basic KA Interactions among Communities of Practice

- *Configuration A: Target community is the focus community.* In this case the primary beneficiaries of the codified knowledge are the practitioners in the focus community. Many communities perform this function through their own research; in this KA configuration, an external (or at least distinct) investigator community brings experience in knowledge organization to such a collaboration. Often the purpose is to create a basis for sharing knowledge across multiple sub-communities within the focus community as a whole. For example, KA can serve to clarify and standardize knowledge within a professional community; e.g., a medical database, designed to allow doctors from one specialty to access new results from a related specialty.

- *Configuration B: Target community is the investigator community.* One common objective of knowledge acquisition is when one community studies another. The knowledge acquisition effort results in increased understanding for members of the investigator community. Academic studies of a work practice are the classic example of this mode of knowledge acquisition; the investigators are the primary stakeholders in acquiring the knowledge (e.g., to complete a dissertation and obtain a degree).

- *Configuration B: Target community is a third, separate community of practice.* In this case, the target community for the workproducts is distinct from both focus and investigator communities; thus, the knowledge acquisition effort, and in particular the investigator community, plays an intermediary, "bridging" role between the focus and target communities. Often, the investigator community includes practitioners with backgrounds from both the focus and target communities, which makes them particularly adept at "bridging the gap." The TCIMS project was organized in this way, with a separate team whose responsibilities were focused on knowledge acquisition.

These configurations do not exhaust the possible KA modes but they provide a starting set for consideration. They also serve as a checklist of the range of stakeholder interests that should be considered in any large KA effort. They can be applied not only to the overall goals of a KA effort but also to the specific audience for any given KA workproduct. In this regard, they are useful in identifying cases where a single workproduct is expected to meet the needs of multiple audiences that may have conflicting representation requirements and interests in the codified knowledge. A common risk in planning a knowledge acquisition effort is to lose track of the intended audience for a particular KA workproduct, or to assume that a single representation will be able to serve multiple needs for multiple audiences.

Configuration A (focus community as target, or audience) will apply to any workproduct that needs to be validated by informants, or by other practitioners within the focus community. However, the same knowledge might also need to be represented in a form that will serve the needs of system builders developing technology for the focus community. Reconciling the needs for codification formats that are amenable to validation, yet still contribute to the formalization of the knowledge, is a central issue in KA.

In addition, the even when the codified knowledge is intended for other target communities the impact of the knowledge on the focus community must be considered. Whenever elicited knowledge is recorded or codified in a way intended to facilitate review and validation of that data by informants in the focus community, that information will be of potential direct interest to those informants, as practitioners, beyond its validation for use by other communities.

Similarly, Configuration B (investigator community as target) highlights the fact that the investigators will often have at least a partial interest in directly acquiring, not merely transferring, knowledge in the areas of focus. Configuration C (distinct target community) emphasizes the fact that, regardless of the immediate target community for a KA effort, the fact that knowledge has been codified renders it more amenable to broader dissemination. This fact can be both an incentive and a barrier or cause for concern, if the knowledge involved is proprietary, classified or otherwise sensitive.

In Section 4.0 we will see how to combine these modes of knowledge acquisition in a plan for a realistic, comprehensive knowledge acquisition effort.

3.2 The Knowledge Acquisition "Canvas"

In the preceding section we introduced the core elements of the knowledge acquisition process and characterized the process with respect to related forms of learning and knowledge creation. In this section we place these elements in the context of the overall Canvas framework which describes their relationships and supports systematic planning of large-scale KA efforts.

The name Canvas is intended to suggest a "weaving" metaphor. Each element of the knowledge acquisition process — investigator, informant, artifact — changes throughout the KA process. The lifecycle of each elements can be viewed as a a ***thread*** weaving through the overall KA process. Each KA session, be it an interaction between an investigator and one or more informants, or a study of an artifact, represents the convergence of several threads. The element tracked by each thread is transformed in some definable ways as a result of the session, and carries these changes forward to the next intersection with other threads in subsequent sessions. The following subsections will describe each of the different types of threads in detail:

- Section 3.2.1 introduces the basic *elements* of the framework;
- Section 3.2.2 shows how the elements combine in a KA *session*; and
- Section 3.2.3 integrates elements and sessions by describing the different *threads* that can be identified for each element, and how each thread unfolds in a series or sequence of sessions.

The resulting framework provides a basis for considering a broad range of knowledge acquisition situations. The framework can be used as an extensive checklist of critical issues to be considered in planning and managing a knowledge acquisition effort. However, not all KA efforts require a process that takes all these elements into account. Section 3.3 both extends and focuses the framework to address the intended scope for this document in more detail. Section 4.0 then introduces the notion of the knowledge acquisition enterprise, a high-level planning infrastructure that assists planners in tracking the progress of various threads through their lifecycle; assessing the impact of specific planning decisions on multiple threads; determining what changes need to be kept track of, and what to expect of a session in which several threads are brought together.

3.2.1 Basic Elements

In this sub-section we introduce the basic elements or "building blocks" of the KA process in more detail: settings, investigators, informants, artifacts and workproducts, and topics. A given KA session involves some combination of investigators, informants and/or artifacts, and workproducts (as session outputs and possibly as knowledge sources). Settings interact with sessions in various ways: as a direct "element" of observation sessions; or as a possible focus on the knowledge to be acquired in other sessions. In addition, settings characterize both artifacts and informants: artifacts studied are typically workproducts created and/or used in the settings of interest; informants are practitioners with experience in current or previous roles in these settings. Topics, more generally, are primary structuring elements for scoping particular knowledge to be acquired. Information is gathered in a session with a specific focus, centered around one or more topics.

Settings

In Section 3.1.1, we introduced the notion of work settings and of communities of practice, and clarified the related but independent nature of these concepts. In terms of the KA-specific community roles, the settings we are primarily concerned with are the settings in the focus community.It is helpful to explore the different kinds of settings that might be the subject of investigation using KA techniques.

The notion of setting implies, but does not require, physical and temporal co-location. A typical setting in the medical domain would be, for example, a doctor's office, night shift at an emergency room, or an accident scene for trauma care. Many KA techniques, such as the elicitation of scenarios as sequences of related tasks or events, are best suited for describing these kinds of settings. To describe an electronic mail discussion of a difficult case history among doctors, spread over a period of several weeks, as a "setting" would be less intuitive.

However, the more central the role of technology (particularly computer technology, distributed systems, etc.) in the work setting, the less dependable is this notion of co-location, either temporally or spatially. For example, consider the work setting of an Internet newsgroup, which could represent a very distinct community of practice linked via a communication channel that supports both global access and highly asynchronous interactions.

With the above proviso, we can consider a setting to be bounded by a scope or frame which includes a location and a performance period, and a set of processes that may be performed by human actors (performers, participants) or executed by automated processes or agents. (We intentionally do not introduce specific terminology for these concepts as they are not essential to the Canvas planning framework, but are more dependent on representations chosen for documenting work setting activities.)

Within a setting during a given performance period, many activities may take place involving many actors or agents. We distinguish as ***practice*** repeated activity which is subject to learning, increased competency and expert status. The human actors can include practitioners from the one or more communities of practice active in the setting, but not all actors need be practitioners in this sense. Thus in the medical setting, patients may be actors or but would not generally be called practitioners.

In the context of technology development efforts, it is conventional to assume that the work settings of interest are those in the focus community, i.e., the eventual operational environments for the systems to be developed. It is also conventional to consider these settings in terms of work flow rather than in terms of existing systems already in place and interacting with practitioners. In a domain engineering context, however, since reusable components to be developed will be deployed in technology *development* settings, these settings may be equally a subject of knowledge acquisition. (By implication, software developers for applications in the domain may be the practitioners studied.)

It is also frequently the case that the domain of interest includes some existing legacy systems. This creates opportunities to study such systems in use, which requires in turn a notion of settings that includes the presence of technology. Such a notion must address not only the interactions of individuals with systems, but also ways in which the systems foster (or hamper) collaborative work among practitioners. A domain engineering context may also necessitate the study of multiple settings, e.g., for different instances of a family of applications. The emphasis on existing systems, at least, can be equally valuable in any technology development effort where the ability to interoperate with legacy systems is a desired objective.

Investigators

Practitioners in the investigator community are called simply ***investigators***; their routine practice includes activities such as interviewing, cataloging results, and studying documents or tapes. The name "investigator" shares with its use in detective stories the skills of digging out information, determining what needs to be asked next, etc. However, unlike fictional detectives, KA investigators are interested in more than information about individual events or persons. A KA investigator is looking for information hidden within the cultural context of the community of practice under

investigation. This means that the KA investigator must have skills of detecting bias, possible ambiguity or misunderstandings of terminology. (Also unlike fictional detectives, KA investigators rarely have to resort to fisticuffs or be able to detect cyanide from its aroma.)

The framework presented implies that the investigator community is as distinct as the focus and target communities. In practice, however, investigators will often be drawn from the other communities involved and may not initially form a distinct community of their own. This is particularly true in cases where the importance of systematic KA, hence the need for specific knowledge acquisition training and skills, has not been recognized. Using an in-house investigator group (i.e., investigators from the focus community) has advantages (e.g., familiarity, access) and disadvantages (e.g., familiarity, shared implicit context and bias with informants). No matter what the background of an individual investigator, however, the nature of the role is to negotiate the knowledge of focus and target communities, utilizing some knowledge of the knowledge acquisition process itself.

Informants

In a knowledge acquisition context, we use the word ***informant*** to refer to any practitioner in the focus community who provides information to the project. In fact, as shown in the example scenarios of Section 3.1.5, in theory any participant or actor in a work setting of interest could be tapped as an informant. We use this term rather than more specific terms such as "user" or "expert" to avoid embedding assumptions about the informant's status within the focus community, of the informant. When planning an interview the term used to refer to the interviewee can impact how they view the process. (This will be treated in detail later, when we discuss general stakeholder issues in knowledge acquisition in Section 4.4.)

While the informant may not be considered (or self-identify as) an expert in the focus community, he or she serves as a knowledge source with respect to some community of practice. By gathering knowledge from multiple informants, we are able to get at knowledge that may not be accessible to a single individual in the community. So, for example, informants might be selected to include many different roles in the community, e.g., administrative personnel, software developers, domain experts, secretaries, support personnel, etc. In cases where multiple settings are being investigated informants from each setting might be consulted.

Artifacts

Artifacts are workproducts created and used in a work setting of interest, or containing information relevant to that setting, that are selected for study as part of knowledge acquisition. Use of the term "artifact" highlights the fact that a sampling of workproducts is made as part of knowledge acquisition; i.e., just as not all practitioners are selected as informants, not all workproducts of the focus community are selected as artifacts. In addition, the ways in which the investigator accesses and analyzes the artifact may be quite different from the ways the same material would be used as a workproduct by practitioners. Since the artifact is a "window" into the work culture, the standards for assessing its quality are also different; for example, it need not be "accurate" in the same sense needed for its role as a workproduct in the setting.

To make confident use of an artifact as a knowledge source the investigator needs to recover information about the context in which the artifact was created (a kind of reverse engineering). This interpretation step in working with artifacts differs from simple gathering and transfer of data. As an example, suppose an equipment maintenance checklist is studied and it is noted that the mechanic has starred certain entries and crossed out others. Some interpretation is required to know what the stars indicate; for example, areas where trouble has been frequent, areas where a breakdown would have dire consequences, or areas that inspectors are most likely to check up on

procedures. Thus an implied association of artifact as "cultural artifact" is not actually far off the mark. The KA process requires attention to the ways in which work culture knowledge and meaning have become embedded in the artifact. The artifact is "workproduct studied in context."

It is also useful to distinguish artifacts derived from workproducts playing a direct role in the focus setting work processes from "interpretive" artifacts that represent reflection or prior codification on the part of practitioners. Since knowledge transfer is a part of the functioning of any community of practice, material such as textbooks, survey articles, training materials, etc., will be important resources to the investigator. However, such artifacts must be distinguished from both "primary" artifacts (e.g., a lab report, an intake sheet, or a system specification), since it is less clear how to place them in a specific work setting context.

Knowledge Acquisition Workproducts

We also distinguish artifacts as "raw data" or "prior art" existing within the work setting independent of the KA enterprise, from ***knowledge acquisition (KA) workproducts*** created as *part of* the KA process. The distinction is important in several respects. Sessions can use as input both artifacts and KA workproducts from previous sessions. But investigators need to know the derivation of each input to interpret its content correctly. For example, investigators must interpret the representation of a system document used as an artifact according to the context of the original developers of the document; however, investigators make strategic choices about which representations to use for KA workproducts that they will produce. Since a frequent objective of codification in KA is refining documented information from less into more formal representations, it may be helpful to include both artifacts and the KA workproducts derived from (or interpreting) them as a single thread.

The distinction between artifacts and KA workproducts can be a subtle one, particularly when workproducts are created in a collaborative fashion with informants, or created with the specific purpose of obtaining validation from informants. For example, suppose an investigator creates a high-level system architecture in conversation with an applications expert and documents this as part of the acquired knowledge from the KA session. Had the system designers chosen to develop such a diagram themselves, it might have been similar to the one produced by the investigator. Nothing in the content or representation format clarifies its status as an artifact or a KA workproduct created by (or with) the investigator's intervention. Only by tracing the process of the document's creation can we make this distinction. Yet the distinction will be critical in maintaining systematic links back to knowledge sources, and in correctly interpreting the material. It is also important to distinguish informant-derived interpretive artifacts (as discussed above) and the codified workproducts created by investigators as an output of a KA session.

Topics

In the Canvas context, the word ***topic*** refers specifically to an area of knowledge held by the focus community that will be the focus of attention in some KA session. Topics are usually aligned with the overall objectives for the KA effort; for example, if KA is being performed to elicit technology requirements then topics would probably involve exploring work processes and products that were good candidates for potential automation. If business process improvement were the main objective, topics might include "breakdowns in work" or "customer complaints."

In the domain engineering context, topics usually fall within the scope of the domain. At various points of discussion in this document, we may use the terms "topic" and "domain" somewhat interchangeably.

When viewed as a way of scoping or filtering a knowledge elicitation session or organizing outputs into separate workproducts, topics can have address both a specific subject area and a particular desired ***knowledge type***; e.g., process knowledge, knowledge of the terminology used, knowledge about the heuristics or rationales underlying decisions, or the context for the knowledge (the work settings that provide the experience base). Types of knowledge and the impact this can have on choice of representations for KA workproducts are discussed in detail in Section 5.0.

It can be useful to structure the various topics of interest into some overall model that shows relationships between various topics. As an example, a skills inventory is often structured in a hierarchical fashion that suggests which skills are components or prerequisites of which other skills. Understanding the dependencies between different topic areas will become critical when planning KA activities in detail.

3.2.2 KA Sessions

Having reviewed the basic elements, or separate strands of the knowledge acquisition canvas, we will now examine the ways in which the various elements coincide in the knowledge acquisition process itself.

A knowledge acquisition ***session*** (KA session) is an event (or set of events) where an investigator consults some knowledge source (a person, a document, an observed process), elicits some knowledge and codifies some of that knowledge into a KA workproduct. For convenience, we treat all these activities as part of the session, even if, for example, the investigator leaves the interview and writes up session notes that evening in her hotel room. It is the full cycle through access of the knowledge source, elicitation, and codification that bounds the session as a whole.

Session Objectives

The planning process determines a set of ***session objectives*** for each knowledge acquisition session, characterized in terms of each component element of the session. Each element is part of a corresponding thread which develops through a sequence of sessions. The desired changes or progress for each element, relative to its thread, constitute the objectives for the session; these must be reconciled in the most appropriate way. The thread history previous to the session creates necessary background or context for the element; it can also create bias, influence of past experience that may compromise the desired results of the session.

Each session, in particular, will have *topic-specific* objectives—the knowledge that is to be elicited in the session. The focus of the session might be directed in various ways: for example, exploring the various kinds of tasks a practitioner performs in a given setting; or tracing the sequence of a given task or procedure in detail. The topics of focus provide a scoping mechanism that helps direct attention to a tractable amount of material to cover in the given session. The type of knowledge to be elicited can also be used to focus how the session progresses; if the session will elicit procedural knowledge, then a different sequence of investigations will happen than if the session is intended to elicit declarative knowledge. In addition, the intended audience for the session results are an important aspect of the objectives. While a knowledge acquisition effort usually imposes an overall intended audience, a specific intended audience can be defined for the workproducts resulting from each session; different workproducts can even be created for different audiences.

The topics, knowledge types and audience choices have some interdependencies, and will help determine the knowledge sources consulted for the session and the format of the session (e.g., one-on-one, joint meeting with multiple informants, walk-through of a document with an expert, etc.), and the representation used for the knowledge acquisition workproduct. For example, if the

topic of focus is a certain set of work activities, then procedural knowledge will likely be the type of knowledge required; certain representations will be more appropriate and certain types of informants will be better sources for this kind of data.

There may be additional objectives defined for any element of the KA "canvas". An investigator's objective for a session might be to gain personal familiarity with a new aspect of the domain as preparation for further sessions. An informant-specific objective might be to obtain knowledge from the informant in a one-on-one interview setting as a follow-up to a group facilitated session. (Note that the informant may have personal objectives or agendas for the session different from the informant-*specific* objectives that are part of the plan.) An artifact-specific objective might be to synthesize information in the artifact with a set of similar artifacts from different work settings, to produce a comparative or summary KA workproduct.

Session Outcomes

The results or outcomes of a given session can also be characterized in terms of their impact on each element that interacts in the session. Each session will have a primary outcome that reflects some degree of having satisfied the objectives. Specifically, some knowledge in the topic areas and of the desired knowledge types should have been acquired and represented in a form appropriate for the intended audience.

Each session will also have unanticipated results that affect some or all the elements. Here, the many forms of learning and knowledge creation related to KA, discussed in Section 3.1.5, serve as a useful reference, since many of these forms may occur as "side-effects" of the core KA process. For example, suppose that during a session in which the investigator planned to elicit the steps taken by the informant in preparing a helicopter for flight, the informant provides information about the various ways in which the helicopter could malfunction. The investigator must be prepared for this situation, and decide whether to stick to the stated goal of the session, or to modify the session objectives to obtain the unanticipated result of documenting some knowledge about helicopter malfunctions.

Unanticipated information obtained might be relevant to the planning of the knowledge acquisition process itself, rather than the topics. For example, an informant might tell the investigator about several other people that should be contacted as potential informants. This information then must be fed back to the planning process in some systematic way. This iterative acquisition of information about knowledge sources usually necessitates a highly adaptive KA planning process.

In addition to the desired and unanticipated knowledge codified as a result of the session, each session has *peripheral* effects as well, such as the creation of new knowledge in the session. We have made clear that KA involves more than just the investigator's individual learning about the topic of focus. Nevertheless, some learning of this kind will take place with every session. The investigator leaves the session with a greater base of experience in the topic area and the work setting. The peripheral knowledge gained during a session forms the basis of what we will call the "thread" of development of an investigator.

Similarly, while the KA process is distinct from the focus community's own knowledge creation activities, informants may learn new things about their own topics of knowledge as a result of the interaction, or may think about their knowledge in new ways. This may occur through the act of reflection required to articulate terms and concepts to the investigator, who lacks some of the implicit context of other practitioners from the informant's community. Or the informant may be brought into contact with other people from the same community that he or she would not have encountered as part of routine work activities; and this encounter may engender new knowledge as well. In any case, the new learning created for the informant is also part of the knowledge elic-

itation event, but not a defining aspect. This descriptive framework directly suggests how the planning process can work with session planning; the session planning process is described in detail in Section 4.9.

A systematic knowledge acquisition process has mechanisms to support explicit documentation of objectives for the session, consideration of the potential impact of the session event and results on the various threads involved, and management of intended and unanticipated results as well as the peripheral "side effects" of each session. In the next sub-section we look at the various types of threads in more detail.

3.2.3 KA Threads

The Canvas framework reflects the essential idea that knowledge acquisition *creates new knowledge* through the elicitation and codification process. Each KA session creates knowledge in ways that support the overall objectives as well as in other, peripheral ways. This implies a sort of "Heisenberg Uncertainty" principle for knowledge acquisition. There is no such thing as passive knowledge acquisition; the process is always an *intervention* of some kind in the work settings of focus and for all participants. In particular, every session has the potential to change:

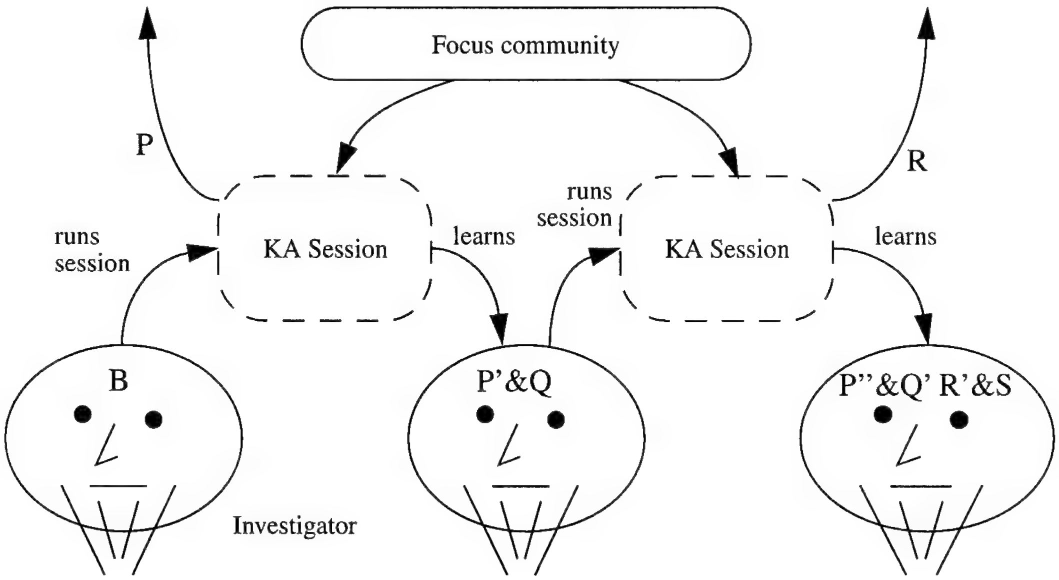
- the knowledge of the investigator;
- the knowledge of the informant; and
- the knowledge codified for an artifact (or workproduct) via the addition of annotations, commentary, or interpretation.

The session can also affect more than just the knowledge of the people involved. Because sessions are interventions in the communities of practice, they may change the role or status of participants as well. To trace the impact of each session we can define separate learning "life cycles" for the various elements of the KA process, including investigators, informants, artifacts and workproducts, and topics. We refer to these cycles as ***threads***.

The paragraphs below give a high-level overview of the major threads that weave through the various sessions conducted as part of knowledge acquisition. A given thread has a focus or subject and traces the "movement" of that subject through a series of sessions, each one of which is influenced by the previous history of the thread for that element, and in turn creates changes that are reflected in the further evolution of the thread. The formal notion is uniform, but the value of the thread concept is quite distinct for each type of KA element. The focus here is on the interaction of the different kinds of threads with particular sessions, as well as some of the overall concerns reflected in the structure of the threads as a whole. More detail on the planning process for each type of thread is provided in Section 4.8.

Investigator Threads

Since knowledge acquisition is a learning and knowledge creation process as well as a transfer process, each session has multiple outcomes with respect to the investigator. As described in Section 3.2.2, there is the knowledge that is codified and made available to the intended audience as well as knowledge that is not codified, some portion of which is nevertheless available to the investigator in later sessions. This means the sequence of sessions that an investigator performs as part of the KA effort will have important consequences for the knowledge obtained and the overall development of the investigator.



Subsequent sessions run by the same investigator; P and R are codified for the audience, but Q and S are also retained.

Exhibit 5. Investigator Thread Interacting with Sessions

Exhibit 5 illustrates the development of an investigator's knowledge through participation in a series of sessions. (Although it shows only the investigator's point of view, a similar type of schematic would be useful for understanding each of the other types of threads as well.) The investigator brings some background knowledge, experience, and possible relationship to the focus community to the initial session (B). As a result of the first session, some knowledge (P) is codified for transfer to the intended audience. But not all the knowledge generated from the session is reflected in the codified workproduct; some residual knowledge (Q) is also retained. Also, strictly speaking, not all of P is retained by the investigator, indicated by P'. A similar process takes place with each subsequent session; some retained knowledge may fade (P'') while new knowledge (R' and S) will be absorbed.

This is, naturally, a simplistic way to view knowledge creation which casts it in overly objectified terms. Nevertheless, it suggests some primary issues faced in the investigator thread, including the following:

- Bias — information known prior to the KA effort and that gained in earlier sessions can interfere with information to be acquired in later sessions.
- Preparation — information gained in earlier sessions might be a pre-requisite for understanding information in later sessions.
- Retention — some, but not all, the information elicited in a session will be internalized by the investigator, thus contributing both desirable background preparation and bias to be considered in subsequent sessions.

  Example. In the TCIMS effort, there were different types of sessions defined as part of the Knowledge Acquisition Plan. The TCIMS plan defined *baseline sessions* to provide a foundation of domain-specific knowledge; they served both to orient investigators to the main technical areas and to orient *domain experts* (the preferred term for the informant role in TCIMS)

to the knowledge acquisition enterprise itself. Later, more *specialized sessions* were conducted as part of the plan. These focused on specific topic areas in more detail, and generally moved through a series of sessions focusing on different aspects (scenarios, tasks, etc.) in succession.

For each investigator, a baseline session ideally preceded more detailed sessions in a given domain or technical area. In this sense, the TCIMS plan included a mechanism for developing an investigator, by starting with participating in or studying the results of a baseline session, and moving on to one or more specialized sessions in a given technical area. This helped ensure adequate preparation, a critical concern in dealing with medical professionals.

In addition to the life cycle regulating each investigator within a given technical area, a given investigator's degree of experience in general knowledge acquisition techniques, in the broader medical domain, or in the use of specific representations was also a factor in selecting investigators for particular sessions. Over time, the plan could assist in the development of the investigator's skills in desired directions.

Informant Threads

Since the KA process creates new knowledge for the informant as well as the investigator, it is important to consider the sequence or series of interactions through which a practitioner becomes involved as an informant on a KA enterprise. (The informant thread notion is not sufficiently distinct from the investigator thread in terms of visual representation to warrant a separate illustration.) There are several critical issues to be considered in the informant thread. A primary issue is making sure that the informant has been "tapped" for the important topics for which he or she has relevant knowledge, with minimum redundancy across sessions. The primary structure of an informant thread, therefore, should show where different sessions have shifted the topic of focus, or, if revisiting a topic, has taken it to a new level of detail or approached it from a different perspective.

> Example. On the TCIMS effort, where medical experts' time was at a premium, an important task of the KA plan was to ensure that the same expert was not asked the same question multiple times. Instead, the knowledge acquisition reports were organized into a dossier that provided a structure so that all investigators could, in principle, access the data derived from previous sessions before scheduling and performing new interviews with a given expert.

In addition, through their participation in sessions, interactions with the investigator, the informant also learn something, because of the intervention of the KA process in their own knowledge state. Interviewing the informant in tandem with others from the same setting will also affect the KA process. The informants' learning lifecycle also forms part of their thread of the KA canvas.

Through interaction with investigators, informants gain increasing familiarity with the goals and approach of the KA enterprise. A certain amount of orientation is required to make someone, even an expert in the field, an effective informant. Many informants will be interviewed only once; however, others may become important continuing resources for the KA project team. In considering the thread of multiple sessions with a given informant, therefore, planners should consider the role of the thread in gradually educating the informant to the knowledge acquisition task. This is almost the dual of the investigator's thread, which represents the investigator gradually gaining more detailed knowledge about the domain.

> Example. On the TCIMS effort, there were a series of specific representations and corresponding types of interviews defined in the plan: e.g., task analysis, scenarios, conceptual analysis, etc. Generally, there was a view that, for a particular informant, there was a most desirable sequence for these types of interviews: i.e., task analysis would be a more accessi-

ble starting point than conceptual analysis. This suggested that, for informants who would be significant knowledge sources interviewed several times, a certain "life cycle" of different types of sessions happening in a pre-defined sequence was an important aspect of the planning process.

Artifact Threads

Planning the life cycle for a specific artifact, as an inanimate information source, is a different problem than planning for the various interactions of an informant.

Clearly, an inanimate data source will not itself be affected by psychological phenomena such as bias or learning.[1] This does not imply that management of bias is insignificant in working with artifacts, which are strongly shaped by the embedded contextual knowledge of their creators and the work settings in which they are created. (In fact, in data gathered from artifacts rather than informants this hidden contextual information may be harder to detect and manage, and often can only be discovered through interactions with informants.) However, this initial embedded bias does not change as a result of subsequent interactions of investigators with the workproduct.

Instead, the thread for an artifact consists of the various phases of *interpretation* performed on it. In Exhibit 6, we see a sample life cycle for a knowledge acquisition artifact: each dashed line in the exhibit refers to the production of new KA workproducts, resulting from a study of the previous one.

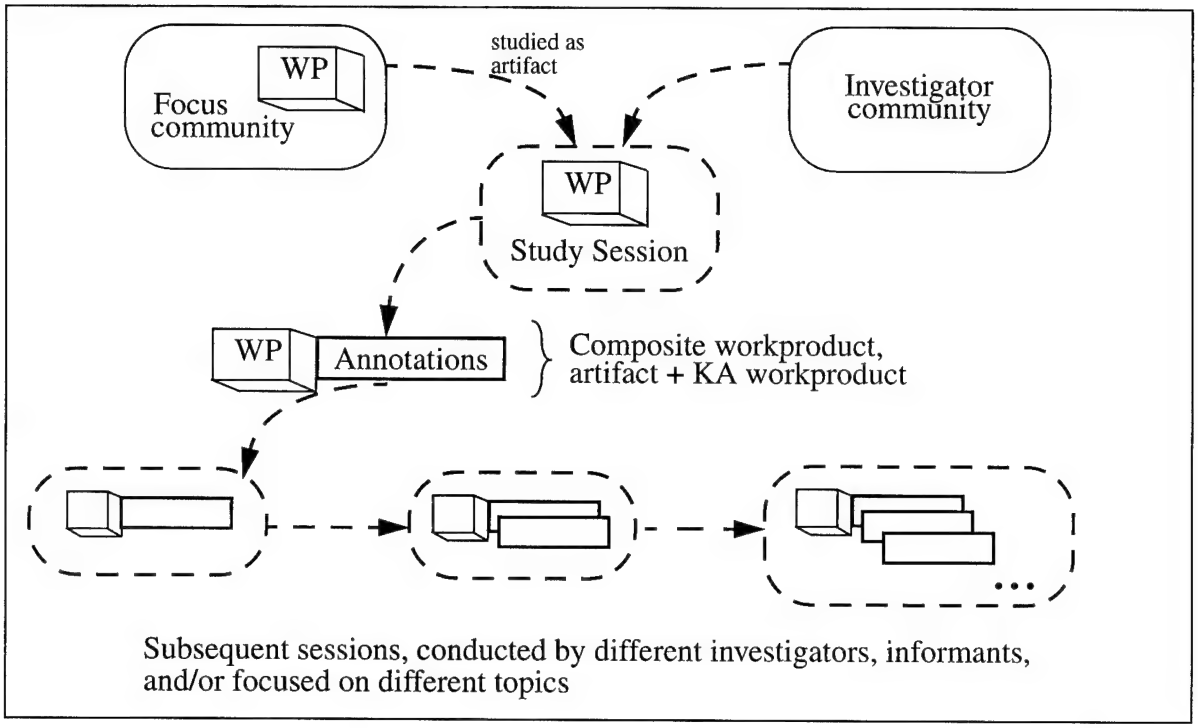


Exhibit 6. Thread of the Life Cycle of an Artifact

[1]. We exclude learning programs or other technologies that suggest the notion of an "intelligent artifact," though clearly these could be accommodated within the Canvas framework via a thread that allowed for the influence of previous sessions on the artifact itself.

The artifact begins its role in the KA effort as a workproduct created by practitioners within the work setting of the focus community. The workproduct is selected as an artifact for study (a knowledge source) by some investigator. Through a study of the artifact, new knowledge is elicited, which is then codified in a KA workproduct such as annotations of features of interest for the artifact. These annotations can be considered part of an ongoing "composite workproduct" that adds a layer of interpretation to the original workproduct. For example, subsequent investigators studying the same *original* workproduct as an artifact could choose to view the annotations of previous study sessions as well. There could be reasons for them to view these annotations or reasons for them to choose *not* to view them (e.g., controlling bias).

The composite workproduct (workproduct + annotation or interpretation) can itself becomes a subject for further study, perhaps by a different investigator. Through such sessions, the workproduct undergoes successive stages of interpretation; some common interpretations are annotation, comparison, and summarizing. This sequence of workproducts, beginning with the original workproduct from the focus community and continuing with successive interpretations added as part of the KA effort, forms the thread for that artifact (or KA workproduct) in the Canvas framework. The KA workproducts in such a thread are called ***derivative workproducts*** of the original focus domain artifact or original KA workproduct.

Issues faced in this life cycle include the following:

- Redundancy — controlling multiple studies of the same artifact.
- Filtering — Determining how to filter the topical material from the artifact as a whole — a particular challenge in domain engineering, where comparative data is desired.
- Dropping through the cracks — Making sure that a given artifact has been examined, or that a decision has been made not to examine it.
- Scoping — Making sure that subsequent interpretations of the artifact fall within the scope of the KA enterprise.
- Proprietary access — Observing restrictions on who controls what can be seen, placed in dossier etc.

We show how these issues interact in the following example:

> Example. In the TCIMS effort, a number of reports were written based on initial interviews with informants. These reports are knowledge acquisition workproducts, and were stored in a repository called SEPWeb. Later in the effort, these reports were retrieved from the repository, and further analysis was done to them; these included summaries and changes of representation. The original reports were considered consortium confidential, and were not released for public viewing, while the derivative workproducts were published on the WWW. Further efforts summarized these workproducts further, to produce models written in a formal modeling language. Along its thread, we see changes in both the level of security and formality of the increasingly derivative workproducts.

KA Threads: Summary

The threads described above — investigator threads, informant threads, and artifact threads — are not independent. Each session potentially plays a role in several threads; i.e., in the life cycles of informants, investigators, and/or artifacts. The sessions described separately in the thread discussions above could be the same session. In our canvas metaphor, each session is therefore a crossing of threads, a meeting point where each thread moves its subject forward along its respective

life cycle. Planning a single session involves simultaneous changes to multiple threads. Conversely, each thread progresses by passing through a sequence of sessions, linked by the common element of the subject for that thread.

The essence of planning knowledge acquisition with Canvas is to manage development of all threads in parallel, according to the issues for each thread type as outlined above. This is a highly reactive planning task, one that requires careful attention to development of all resources of the project. The outputs of a given session that may affect subsequent decisions cannot be known in advance; planning must therefore be iterative and some effective means of communication of results across sessions, among investigators must be provided.

3.3 Issues Implied by the Framework

The core concepts described earlier in this section provide a useful framework for a wide variety of KA applications. The emphasis of this guidebook is on two particular kinds of KA enterprises, typified by applications to date of SEP and ODM.

Traditional approaches to KA have limitations deriving from the evolution of these approaches from work in the social sciences or linguistics, or in knowledge engineering for expert system development. The necessary extensions include the following aspects:

- Knowledge acquisition inherently involves cultural communication.
- Performing knowledge acquisition within technology-oriented settings challenges certain traditional KA terminology and assumptions and requires new techniques.
- Knowledge acquisition needs to address explicit management of variability in the data.
- Knowledge acquisition, at its best, is a collaborative process.
- The knowledge acquisition process intervenes in the communities where it is applied.

Each of these five aspects is discussed in a separate sub-section below.

3.3.1 KA as Cultural Communication

According to the definitions provided in Section 3.1, knowledge acquisition most often implies transfer of knowledge across distinct communities of practice. Since in each respective setting knowledge may be expressed in varying terminology and put into practice in very different ways, KA inherently involves a cultural shift which can be viewed as a highly specialized form of "cross-cultural communication." A well-known example of this challenge comes from the history of expert systems development, where knowledge embedded in the medical community had to be made available to a community of program developers and AI researchers in order to produce medical expert systems.

The element of cultural communication can come into play even when the primary role of the investigator is to facilitate knowledge exchange within the focus community, since communities of practice are not homogenous knowledge environments. For example, within the medical community itself, there can be cultural differences, fostered in part by organizational dynamics; most hospitals today are organized in such a way that there is considerable cultural difference between doctors and nurses.

It is well-known that translation from one professional or work setting to another is fraught with potential for misunderstanding and misinterpretation. Since we are often prisoners of our own cultural perspective, we are unaware that this cultural translation and re-interpretation process is taking place. However, knowledge acquisition activities often take place in the context of projects where these dynamics are not visible. Software development is not often viewed from a cultural communication standpoint. Even in the social sciences, where understanding of cultural dynamics is a given, there have been few systematic studies of the knowledge acquisition process itself — i.e., the cultural act of studying another culture—as a form of cultural communication [14].

Canvas provides a framework for anticipating and planning for both the potential benefits and the risks of KA as cultural communication. By stressing the central role of codification and transfer, the potential influence of bias, and the influence of distinct (focus, investigator and target) community roles, Canvas enables characterization of the knowledge transfer challenges in a given KA situation with some precision, as an aid to systematic planning.

3.3.2 KA in Technology-Intensive Settings

Although KA has been utilized in the service of technology development, there is still a lack of good methods for doing knowledge acquisition *about* technology use. Some of the reason for this is historical.

- Within social science research, knowledge acquisition techniques evolved from studies of non-industrialized cultures and were only gradually applied to the culture of the researchers themselves (i.e., anthropology to sociology).
- Many methods for data or knowledge acquisition reflect a technology development context where the primary goal is automation of manual processes. In the era when the first major computer-based systems were put in place, most business work environments were based on manually oriented work processes, with information exchanged via paper, face-to-face meetings, conversations, etc.
- Even early expert systems development tended to be initiated in domains where little automation was present; this was the goal was to automate decision processes considered too intractable for contemporary software engineering techniques. Thus the problems of integrating knowledge-based systems with legacy software systems only gradually came to the fore.

However, in current conditions it is a reasonable default assumption that new or evolving systems will be deployed in a technology-rich landscape, with numerous existing legacy systems in various stages of implementation, usage or decay. Even opportunities for expert system development now occur more and more frequently at the frontier of established system capabilities. This creates challenges as well as opportunities: aspects of technology-intensive settings have ramifications throughout the KA process. KA techniques must be adapted to shed light on how work practices evolve in relation to and are influenced by computer-based systems within the work setting. KA planning must account for the influence of legacy systems in the setting and in the experience of practitioners, as well as the influence of trends, business requirements and expectations about changing system capabilities.

Challenges and opportunities of doing KA in a technology-intensive setting include the following:

- Technology is not created in a vacuum. It is created within a technologists' community of practice that has complex interactions with the user community where the technology is eventually fielded and used. If we want to study a technology-intensive environment and pay attention to the technology aspects, we must gather data about both these communities.

This is not a simple undertaking, however. Technologist and user settings generally represent quite distinct communities of practice. As mentioned above, the problem of "cross-cultural communication" arises in the cultural clash between developer and user communities. These communities "collide" in the settings where the technology is used, as well as in the process of gathering requirements; to the extent that these requirements gathering must reflect knowledge about the user's environment and work practice, it constitutes a form of KA.

A familiar example is the problem of software error messages. The error categories and terms are often determined and described by the developers, but inscrutable to users. This is not simply a problem of poorly describing the errors; it is a problem of distinguishing and mapping categories of errors important for the two groups. KA that would serve a technology development goal has the task of examining the already existing legacy as "cross-cultural communication and conflict, embedded in software artifacts".

- Trying to investigate software developers as informants also challenges some typical presuppositions of KA. Artifacts may be highly formal system design documents requiring specialized knowledge to interpret (not different in principle from detailed medical reports that require a specialists' interpretation). The work practice of software developers is not easily amenable to description via workflow-oriented task and scenario descriptions. Recognition of KA as a distinct discipline is not prevalent in the software engineering field as a whole. Therefore, it is likely that someone with a technology background will be "drafted" into the investigator role on a technology-oriented KA effort.

- It can be harder to gather data about work practices mediated by technology using traditional techniques such as interviewing and observation. Since technology externalizes many routine tasks, expert users may not be able to explain their procedures out of the context of the technology itself. Significant processes may take place with the flick of a key at a keyboard and hence be nearly invisible even to direct observation.

  On the other hand, the presence of technology in the workplace also creates opportunities for knowledge acquisition. Investigators can experiment directly with legacy software as part of knowledge acquisition. Through instrumentation automated logging and collection of usage data can be performed in a relatively non-obtrusive way. Since every fielded system will need to evolve over time in response to needs and limitations discovered by users, some consideration for "system as knowledge acquisition aid" should be a routine part of the technology design process.

- If KA is being performed to provide information to technology developers, a special set of stakeholder issues may arise around expectations about and resistance to the introduction of new technology. Even where there are no clear requirements for new applications, technologists' own product ideas will exert influence on the data deemed valuable and the kinds of questions that are asked. Similarly, users have ideas about desired system functionality, or resistance to introduction of technology, which affects the information they offer.

Knowledge acquisition is deeply woven into the entire process of system specification and development; in fact, some form of knowledge acquisition must precede almost any requirements analysis phase. The kinds of issues dealt with in the KA enterprise should be of paramount concern to those creating technological systems for use in people's everyday work practice.

3.3.3 Systematic Treatment of Variability

Variability has not always been an element of explicit focus in conventional approaches to KA. It can refer simply to "noise" in the data. It can refer to lack of consensus in expert opinion. But explicit documentation of variability is important for certain KA goals; in particular, accommodation of variability is of prime importance in domain engineering.

> Example. If three medics describe a standard procedure and give three contrasting sequences of tasks, are two of them wrong? Are they describing particular events, "routinized" generic descriptions, or the official textbook version of what's supposed to take place in the given situation? Are there different practices involved because the three medics work in different units, different hospitals, different states?

Historically, knowledge-based or expert systems have been oriented towards domains of high-level, professional knowledge. Variability in opinion or belief may be interpreted as a sign of instability in such disciplines. Hence it is to be expected that informants might be reluctant to emphasize variability aspects, and/or that interviewers might simplify data to emphasize a common or generic result.

In previous applications of SEP, variability was dealt with primarily in the later, architectural aspects of system development. Variability was not addressed explicitly in the KA workproducts themselves. In the KA phase, investigators sought the common view of the experts they studied, which sometimes involved reduction of variability. How this common view emerges out of data with inherent variability is one of the points where the "magic happens" in a typical KA process.

In contrast, domain engineering for reuse, as exemplified by the ODM process, requires a specific search for patterns of variability. For the purposes of reuse, variability will generally surface in comparisons across similar systems within a specified domain. A specific "representative set" of knowledge sources must be selected that provides sufficient data on the variability to be modeled. The sample set of data selected and elicitation techniques employed are designed to explicitly *include* rather than reduce variability. In the reuse context, this variability data helps to define an intended multi-use scope of applicability for components to be developed.

Although domain engineering highlights the need for systematic treatment of variability, the general concept of variability has rich implications for a KA framework applicable outside a formal domain engineering context. There are valid reasons to pay closer attention to variability issues in any KA effort. We can distinguish "various" kinds of variability that may need to be addressed in knowledge acquisition. Some are relevant to many KA tasks, others are more specific to a domain engineering context.

- *Variability within a given work practice setting.* Knowledge-intensive activity in any work setting typically involves making expert judgments among similar cases or conditions. This variability is the routine variability encountered in work practice; it needs to be negotiated by performers in that setting in order to perform competent or expert decision-making and judgment. An expert can recognize and classify cases, deal with ambiguity, and deal with "outliers" that somehow violate the assumed scope factored into the classification.

  Example. Doctors in trauma care settings will encounter patients with a wide variety of trauma conditions to be treated, in conditions where there are insufficient resources (beds, supplies, medical staff available). Triage procedures in such an environment require being able to discriminate between a wide variety of specific injuries and conditions, and reduce these rapidly to a few categories relevant for immediate treatment decisions (e.g., lost cause, can survive without immediate attention, or immediate attention could make a difference).

A single expert's opinion thus represents a certain model or organization of a range of variability. This kind of variability can be represented with semantic models (concept models, taxonomies) that capture expert practitioner knowledge.

- *Variability across work settings.* When multiple work settings are studied, variability can be observed across these settings. These variations may reflect differences in language or terminology, or variability in practices and procedures. Many representation notations for describing activities within a single setting provide ways of capturing alternative decision paths, etc., but not of reflecting diverse settings. These differences also might not be available as an established "model" from some informant, but might emerge as a composite of gathering analogous data in multiple settings, through a collaboration between the investigators and informants.

  Example. Continuing the triage example, triage practices might differ from doctor to doctor; there may be different policies in place at different medical institutions or in different countries, etc. These differences could provide valuable data in knowledge acquisition. Certain informants who had the benefit of experience in multiple settings might have greater awareness of these differences and be able to comment on them. Changes in practice over time within a given setting may also be a source of data, as observed by people with longer history of work practice within the setting.

- *Variability in system requirements.* An implemented system always reflects the developers' implicit or explicit model of users' work practice. This model will reflect certain dimensions of variability, in which case the system will be designed to respond or adapt to this variability. Conventional systems development, targeted towards building systems to support work in a given setting, has focused more on automating common, routine procedures, leaving performers free to focus more on knowledge-intensive tasks. Thus variability representations in conventional systems design focus on dynamic "procedural" models such as iteration, selection of alternatives, partitioning into process and data, etc. Expert systems might attempt to capture more of the variability that is a factor in expert performance. In any case, once the system is implemented, it provides a different source of data than the practitioner community.

  Example. Consider a computer system that logs the intake record and medical case histories of trauma care victims in emergency situations. Because of issues of liability it is desirable to record more than just the treatment administered to individual patients, since viewed in isolation decisions made following a triage protocol might appear to have been improper treatment. The computer system will represent an "abstraction" of the triage process that might embed assumptions of the process, the various types of rationale used in decisions, etc.

  The application developers have a choice: they might provide a hard-coded checklist of diagnosis codes, a checklist tailorable as part of system customization per installed site, a checklist that can be arbitrarily extended by the user, or even a free text field. This design decision, which yields a system feature in the implemented system, reflects a "theory" held by the developers about where variabilities in practice might be introduced (i.e., different diagnosis codes), the range of that variability, and its desirability. For example, new diagnosis codes could be entered over time, but variability in the form of misspellings or inconsistent usage of diagnosis codes would be counter-productive and guarded against.

- *Variability across systems.* Domain engineering for the purposes of designing reusable software components must deal with variability at several levels simultaneously. The end goal is to create components or other assets that will be utilized by developers building software systems for a given domain. A domain-specific language is formalized to describe the semantics of particular components in terms of the domain-specific functionality they provide. However, since these components are intended to be useful in building many systems, this language must be expressive enough to cover the anticipated variability across those intended

target systems. Each potential "application context," a system development effort in which the components could be used, will have been built (or will be built) with a specific usage setting in mind. The domain model must be able to express variability in those multiple usage settings to the extent that they affect the components themselves.

This variability can be studied in at least two complementary ways. The first is through comparative analysis of multiple usage settings, or multiple system artifacts created for those settings. The second is by studying the development setting directly, particularly the processes performed by system developers building applications in the domain. It may also be necessary to collect information on a given system from the standpoint of both the processes by which it is created and the system products themselves.

Elicitation of variability information can pose particular challenges in the planning of a knowledge acquisition enterprise. The implications of a concern with systematic management of variability for planning the KA process is discussed further in Section 4.9.6.

3.3.4 Collaborative Knowledge Acquisition

A careful reading of the definition of knowledge acquisition in Section 3.1.2 reveals that there is no condition that the knowledge that is elicited, codified or transferred to a new community be "held" in any way by the informant. This is an intentional omission, which has considerable impact on how we view the knowledge acquisition process. In particular, although knowledge acquisition can include a simple transfer of known information from informant to investigator, the most profitable cases of knowledge acquisition are those in which *new knowledge is create*d as a result of the collaboration between investigator and informant. A number of Canvas features provide specific support for this view of collaborative creation of knowledge, from the inclusion in the planning process of a thread specifically for tracking and managing the informant lifecycle, to the treatment of representations as aids for knowledge acquisition.

The process by which interaction between investigator and informant creates useful knowledge that was not available to either of them alone, is called ***collaborative knowledge creation*** or simply ***collaboration***. The Canvas framework has been designed to treat collaboration as the default mode of knowledge acquisition, rather than as an unanticipated side-effect or special case.

Example. A common complaint leveled against the field of expert systems development is that many of the systems are used only as tutorial systems to train new experts (e.g., [7]). The original motivation for most expert systems projects was that in some fields, certain knowledge was held by a small number of highly specialized experts, who would someday change jobs or retire; the organizations for whom these individuals worked wanted to capture this expertise as an asset that would last beyond the tenure of the individuals themselves. The expertise had been gained through many years of experience, usually with machinery or systems that were not well understood at the time, and hence was open-ended and full of mysterious items shrouded in the history of the system. No training materials were on hand to transfer this expertise to new practitioners by training, and the cost of apprenticeship was difficult to assess.

The process of constructing the expert system included a phase of knowledge acquisition, in which this knowledge was organized, regularities were noticed and documented, and the resulting body of knowledge was then codified, often in the form of a rule-based system. This process of organizing and documenting the expertise removed much of the motivation for having an expert system; in particular, the expertise was no longer open-ended and mysterious, and it was now reasonable to imagine teaching it to a novice, or even starting a conventional software development project to support it. The rule-based system, which was intended

to perform the expert's task in his eventual absence, was deemed a failure (since it never took over anyone's work). Instead, it took over the role of training material for new practitioners.

The moral of this story, as it relates to the Canvas approach, is that the collaborative knowledge creation dynamics of the KA process were somewhat invisible to the technologists—focused as they were on producing the requisite expert system—but became quite evident to the experts themselves. The most important value was obtained by helping the experts codify their knowledge in a new way.

Limiting Assumptions

As the example above suggests, the collaborative view challenges some assumptions about knowledge and knowledge acquisition that, while not often stated explicitly, have influenced previous approaches in the field. Here we list a few of these assumptions, along with the contrasting point of view supported by the Canvas framework.

All knowledge can be codified and transferred.

A background of formal education tends to make us think about knowledge as some sort of "material" that can be put into a book, and gained by reading it. There is something comforting about being able to hold up a book and say, "I know what is in this book," and being able to write down "everything I know" about some subject. The assumption is that everything we know should be codifiable in some way in the printed word.

An extreme contrast to this viewpoint is the belief that no knowledge can ever be codified, and that anything that one really learns comes from experience. In certain professions or groups of practitioners this belief can be a strong factor to consider; it may be in part a reaction to attempts at codification that did not allow for the value of informal and undocumented knowledge.

The Canvas approach takes a middle ground; accepting the fact that the result of any learning experience can only be partially codified, Canvas nevertheless relies on the assumption that *something* can be codified, so that it can be recovered by someone else at a later date. When that someone else comes from a different work setting, as is the case in knowledge acquisition, then the problem is to write something that can bridge the gap. The Canvas framework helps to formalize codification as distinct from individual learning by making the target audience for the codified workproduct explicit and emphasizing that this audience is within a distinct community of practice. At the same time, through the concept of threads, the framework also allows the planner to account for the role of uncodified knowledge as both an advantage and a source of bias that needs to be controlled.

Knowledge is held by a single person.

A simple view of knowledge is that it is something "held" by a single person; during knowledge acquisition, we will "mine" this person for the knowledge, which we will store somewhere. This view of knowledge is consistent with over-reliance on interviews with a single expert; the interview is treated simply as a transfer process. This view is also suggested in the metaphor of a "knowledge acquisition bottleneck," where some finite stuff called "knowledge" has to be poured through a limited channel.

In contrast, in the Canvas view, knowledge comes from some interaction, either within the work setting of the informant itself, or in the interaction with the investigator. There is, of course, something very special about the informant chosen for a knowledge acquisition project; we could not take any person off the street, to use as an informant for the domain we are exploring. However,

the result of a collaborative knowledge elicitation session will include things not known to that individual at the start of the session. These might be beliefs internalized in the informant's work culture, or previously unremarked correlations among simpler facts of which the informant was already aware.

The Canvas view on this issue is very similar to that espoused in Winograd and Flores in [59], with the exception that Canvas examines its implications for knowledge acquisition. In particular, because of the variety of ways in which knowledge can be embedded in a work culture, Canvas encourages consideration of a variety of elicitation modes (group sessions, walk-throughs, reviews, solo interview sessions, ethnographic interviews, observation etc.) to create conditions for making tacit knowledge explicit and accessible.

Practitioners don't use models.

In knowledge acquisition, some practitioner from the source setting is selected and used as an informant about some topic in that setting. Knowledge representation typically involves some sort of modeling of this practitioner's knowledge. A simple intuition says that in his own work practice, the practitioner does not use models of his own behavior; he simply acts. Thus the modeling is the responsibility of the investigator or subsequent modeling experts.

In contrast, the Canvas approach is based on the recognition that while certain aspects of an informant's work practice are inaccessible to description and introspection, many informants nevertheless do construct models of their behavior. In some cases, like professional medical practice, these models are quite thoroughly worked out and are very sophisticated. In cases where software developers might themselves be informants, they may be quite familiar with the same representations as those used by the investigators. Canvas takes the view that there are different types of models for different purposes; a careful choice of model can exploit the informant's own modeling experience to encourage collaborative creation of knowledge.

Investigators reflect an informant's knowledge back to him.

Related to the mining view of knowledge acquisition is the view that an investigator acts as a mirror that reflects the informant's thoughts. The informant is not reflective in the sense of thinking about what he does; he simply acts and recounts, and the investigator does all the reflection. This is also related to the idea that informants do not make use of models. This view incorporates two assumptions; namely, that the informant is unreflective, and that the investigator is "neutral."

In contrast, in Canvas, the reflective expert is acknowledged and valued; the role of the investigator is more collaborative, and neutrality is not assumed. The investigator participates in the reflection process of the informant by providing advice about representations, organizing principles, models, or whatever is needed to organize the investigator's reflective process.

Investigation can occur without intervention.

The idea that one can study human behavior without interfering in it has attracted a number of supporters from widely varying fields. Anthropologists have attempted to perform field work that does not intervene in the culture they are studying, while logical positivists view expertise as a set of rules, which can be taken from their context and used independently.

In contrast, Canvas is based on the premise that the most valuable and meaningful knowledge acquisition activity *does* have an impact on informants and their work practice. The new, collaboratively created organization of knowledge will change how the informant thinks, and how he responds in new knowledge acquisition sessions.

A Collaborative Knowledge Creation Scenario

Having debunked these various strawman positions, we can now offer a basic scenario for a collaborative modeling session. In such a session, the investigator may play as much the role of facilitator and coach as interviewer and knowledge gatherer. The informant is encouraged to share responsibility for how the knowledge is codified and how that process and representation serves the needs of the focus community the informant represents. The session is viewed as an opportunity to reflect on knowledge of the work setting and capture it in new and innovative ways. Some specific aspects of such a session might differ from a less collaborative approach:

- The investigator does not take a neutral stance, although allowing for the influence of bias in the session. Relevant knowledge from related domains may be offered as a technique for sparking new insights—not, though, for imposing a normative view on the informant or forcing changes of terminology upon them.
- Informant and investigator can work together on the question of finding appropriate representations for the knowledge. These could use representations from the focus community, documented so that they are understandable to outsiders; or could involve representations offered by the investigator, perhaps with some training of the informant as part of the session.
- This allows the session to be much more representation-driven; the creation of shared, formal workproducts can become one organizing element for the session process (though not the only one, preferably). As a result the amount of time required to "debrief" and filter the information from a session could be greatly reduced, as would be the need for separate validation processes.
- Ideally, the processes of elicitation, codification, validation, and transfer could be integrated into much more flexible cycles, which could take place within the session itself, over the course of several sessions, or over a longer cycle; and involving various participants in the focus, investigator and target communities as appropriate.

3.3.5 KA as Organizational Intervention

In the previous sub-section we introduced the notion of a collaborative knowledge acquisition session where intervention in the informants' knowledge and view of their work is an accepted element of the process. In this sub-section we take this idea one step further, to consider the broader impact of the KA process as a whole on the organizations involved.

KA is presented here as a form of applied or action research that has exceptional capacity to appropriately intervene in a number of ways in the organization and management of work and workplace culture. From the Canvas perspective, such "interventions" always happen whether by intention or not; the goal is a planning framework that allows the interventions to become as intentional as possible and permits their evaluation in terms of technical, organizational goals and even ethical considerations.

KA practitioners need to recognize that the research process itself can influence and intervene into the culture of study. As outlined above, merely making knowledge and practice explicit or asking a new question can change how an informant views his workplace. Such intervention can be positive or negative, depending on the purpose. The key is self-awareness and intentionality of the researcher and goal clarity of the research effort. Being too invasive by imposing unexamined cultural biases is always to be avoided.

A wonderful irony regarding this point is the following. "Pure" ethnographers make it a point of pride and ethics not to intervene in the culture under study. Yet, to some extent they always do, usually in unseen ways. Consultants and managers, on the other hand, are forever looking for new and better ways to intervene and change work cultures; yet they are constantly frustrated by the difficulties of making change. (Perhaps coincidentally, they are often better at offering advice than gathering data; could this have something to do with receptivity and resistance of the client communities?) So, the ones trying not to make changes can perhaps stir up more change than the ones struggling to get change to happen. Go figure!

The KA process can serve as an intentional organization intervention in a number of ways, as described in the following paragraphs.

Facilitate new conversations.

The LIBRA approach to assessment for reuse [48] stresses that organizations change by engaging in new conversations — new topics, new players, new ways of engaging players. The KA effort can bring people together who never knew about each other or never talked to each other or who held a simplistic view of each other and their work. So it can serve to build bridges even with a given setting around the planning of the project, selection of informants, etc. KA sessions themselves are new conversations, and the resulting report or model can be used within the target or focus communities for enabling still more new conversations. New conversations create shared language, understanding, motivation, and possibly ground for future change.

Hold up the mirror about work practices and contextual influences.

The KA effort focuses attention with the combined expertise of the investigator, the informant, and the connection between them provided by the organization and its management. The result is a greater capacity to examine what the work practice and the knowledge being used actually is. This can lead to clearer understanding of the need for change as well as the need to stay the same.

It can also show in dramatic relief how such organization factors (e.g., business structure, reward systems, training, job posting policies, health policies) stifle or enable day-to-day activities within a given community of practice. This data can be of value to leaders and policy makers who are unaware of (or insist upon not recognizing) the impact of global decisions on local action.

Bring attention to work processes, customers, and collegial relations.

In the 1990s, attention to organization structure and process has shifted significantly away from vertical and function units that do not directly serve customers to lateral, cross-functional business processes that are customer focused. In a work environment where there is increasing pressure to think about work and management by lateral process, work process discoveries accessible to a community of practice-based KA approach could be invaluable. Codifying knowledge in action is a significant activity that pictures the work process as a series of steps or deliberations in the service of some customer value, by a network of practitioners in collegial relation to each other. KA efforts may serve to model these key elements of organizational and process redesign via the knowledge codification process.

Bridge different work cultures, especially technologists and users.

A persistent problem in software-intensive user communities and software development communities is their difficulty in communication across the boundaries of their different languages, goals and customer pressures. Via collaborative modeling in the KA effort to discover new knowledge both in KA practitioners' collaborations with technologists and separately with system users,

more grounded and objective ways of describing each community is possible. Where the possibility of greater understanding and connection between user and technology cultures is possible as a result of the outcomes of the KA effort, even greater synergy could be achieved.

A conscious KA effort to bridge the two communities might take a further step than collaborative KA modeling between informant and investigator. It might facilitate the collaborative modeling of knowledge that spans both settings by arranging interviews or focus groups so that user and technologist are essentially doing KA with each other under the guidance of the outside investigator. Clearly, this form of KA would also facilitate intervention among stakeholders using the knowledge creation modeling and codification process to shift major entrenched elements of workplace culture. In the long run, we believe this could be the most significant contribution of an approach to knowledge acquisition based on the core concepts described in this section.

4.0 Knowledge Acquisition Planning

This section builds on core concepts presented in Section 3.0 to present more specific guidance on developing a plan for the knowledge acquisition (KA) enterprise. The section discusses major decisions and issues to consider at various levels of planning, i.e., at the level of the KA enterprise as a whole, phases of KA within the enterprise, specific "threads" of the KA canvas, and individual KA sessions. These are presented in a series of planning activities or steps that guide creation of a workproduct called the ***knowledge acquisition plan*** (or ***KA plan***). The plan addresses the varied aspects of knowledge acquisition planning, including establishing the overall goals, making specific selection of resources (investigators, knowledge sources), and managing the ongoing process. The steps outlined in this section will serve to reduce many risks associated with KA efforts. The framework is most useful by raising certain questions and issues that can help to set realistic expectations for the KA effort at the outset.

This section is intended for use by people who are faced with the *actual task* of planning for a large-scale knowledge acquisition enterprise. It includes detailed and fairly prescriptive advice. If you are reading through the document primarily to glean the main concepts, reading the initial sub-section (Section 4.1) gives a reasonable overview of the planning process. It may be advisable to postpone detailed study of the remainder of the section until a concrete planning task is at hand.

Exhibit 7 indicates the three primary elements of the KA planning and management infrastructure. The first element, the KA plan, is the focus of this section. The dossier (discussed in more detail in Section 6.0) contains links to all the artifacts studied as part of the effort, and all workproducts created by investigators, together with supporting contextual information.

The KA team guidelines provide a structure for standardizing the procedures to be followed throughout the KA effort. A central element of these guidelines is selecting the repertoire of KA representations that will be used for the effort. These are discussed in more detail in Section 5.0.

There are many cross-connections between the elements of the KA planning infrastructure, at various levels. Ideally, the KA plan will be linked into the dossier and the guidelines document. Since adaptive re-planning will be necessary, these connections will also need to be modified and re-validated on an ongoing basis. For example, lessons learned about working with representations can be captured in the guidelines, which in turn help determine further phases of planning. Clearly, automated support would greatly facilitate the planning approach suggested here. Even in an environment without automated support, it will be helpful to envision the KA plan as an active coordination resource and scheduling process rather than merely a static document.

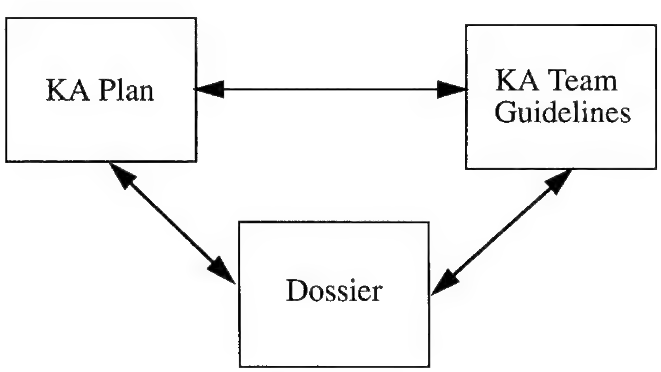


Exhibit 7. KA Planning and Management Infrastructure

Benefits of KA Planning

Many of the activities managed by a coordinated KA plan appear non-technical in nature (e.g., interviewing or reading documents as opposed to generating detailed designs for a software product). Many of the true costs of KA may stay hidden (e.g., time spent with informants). For these very reasons such efforts can consume a lot of resources and can expand unpredictably. Therefore, while it may be tempting to dismiss the need for careful planning, for any project of significant size the potential benefits of properly planning and managing the KA effort are considerable.

More importantly, the planning process might make the difference between success vs. failure in the effective transfer of knowledge acquired to the intended target community. Merely gathering a large volume of high-quality data will not ensure successful transfer, and in fact, given the influence of other factors (addressed by the planning process), might even work against a successful outcome. For example, if extensive resources are expended with unclear criteria for successful outputs, the entire KA effort might be labelled as unfocused or irrelevant, and even useful information may not get used. Furthermore, without a plan that specifically addresses issues such as management of bias, sampling strategies, etc., it can be difficult to evaluate the quality or relevance of the knowledge acquired.

Developing the KA Plan

We have elected here to provide a workproduct outline rather than a detailed formal process for KA planning.[1] The KA plan outline simplifies a process that may involve many collaborative decisions and ongoing revisions. The resulting structure is a useful way of organizing the results of these many interdependent planning decisions. While the sections also offer a reasonable starting point for sequencing planning activities, there will be some iteration required, and it may be advisable at times to jump forward to later sections of the plan to provide tighter focus for some activities. In addition to presenting an overall outline of the KA plan and corresponding process, this section offers more detailed suggestions for certain steps of the process. These include checklists of questions to consider or relevant characteristics to note, and suggested formats for certain planning steps (e.g., knowledge community configurations).

The recommended structure of the KA plan is shown in Exhibit 8. The sections below provide specific guidance for the material for part of the plan. Each KA effort will tailor this plan in ways appropriate to their needs and resources.

The amount of detail in the sections to follow may give the impression that KA planning is itself a major investment. However, the initial planning process need not require a great deal of time and effort. The actual planning activities involved could take place in a few key meetings with primary stakeholders, an investment that should more than pay for itself. Sections of the KA plan other than thread and session plans can be developed prior to commencing detailed knowledge acquisition activities and then re-addressed when necessary as knowledge is gained. The initial KA plan should be validated by project sponsors to ensure that the plan has captured their intentions and that expectations are set reasonably. Once the plan is complete and validated and the infrastructure is in place, KA activities can be initiated.

The need for systematic KA planning may "sneak up" on participants after some informal KA activities have already taken place. This should not discourage the establishment of a more formal plan at the point where its value becomes apparent. Threads for the various sessions that have

---

1. A full process model does not make sense for KA planning in isolation from the KA activities themselves, but detailed guidance for the latter are outside the scope of this document.

1) Enterprise Context (Discussed in Section 4.2)
   - KA enterprise charter
   - KA enterprise strategic stakeholders
   - KA enterprise scope
   - KA enterprise constraints
2) KA enterprise objectives (Discussed in Section 4.3)
   - Community of practice view
   - Knowledge transfer configurations
   - Specific objectives
   - KA Phase Plan
3) KA stakeholder interests (Discussed in Section 4.4)
4) KA elements selection (Discussed in Section 4.5)
   - Settings characteristics
   - Investigators characteristics
   - Informants characteristics
   - Artifacts characteristics
   - Audience characteristics
   - Topics characteristics
5) Representation assignment (Discussed in Section 4.6)
6) Dossier infrastructure (Discussed in Section 4.7)
7) Thread plans (Discussed in Section 4.8)
8) Session plans (Discussed in Section 4.9)
9) Capability Development Plan (Discussed in Section 4.10)

Exhibit 8. Outline of a Knowledge Acquisition Plan

already taken place can be used to initialize the plan; and information gathered prior to the start of the planning process should be documented in the initial dossier,

The KA plan (especially thread plans and session plans) will evolve over the life of the KA effort and should be updated as necessary. Initial plans will be developed for each thread and session respectively. Since each session may intersect with multiple threads, and each thread unfolds through multiple sessions, sessions and threads are closely related and must often be planned jointly. Thread and session plan development will be ongoing during knowledge acquisition, since

all sessions cannot usually be scheduled at the start of the KA effort, and knowledge gained in one session may allow better scheduling of future sessions. As KA activities are conducted, the results are documented in the dossier, and objectives and plans for new KA sessions are iteratively adjusted in dynamic response to interim results. Detailed guidance in performing the KA activities themselves (e.g., conducting an interview, observing work practice, studying a system artifact as data) is beyond the scope of this guidebook. Exhibit 9 shows a schematic illustration of the relations between the different levels of the KA plan. (Note: the figure is intended as an illustration only, not as a suggested notation to use in the planning document.)

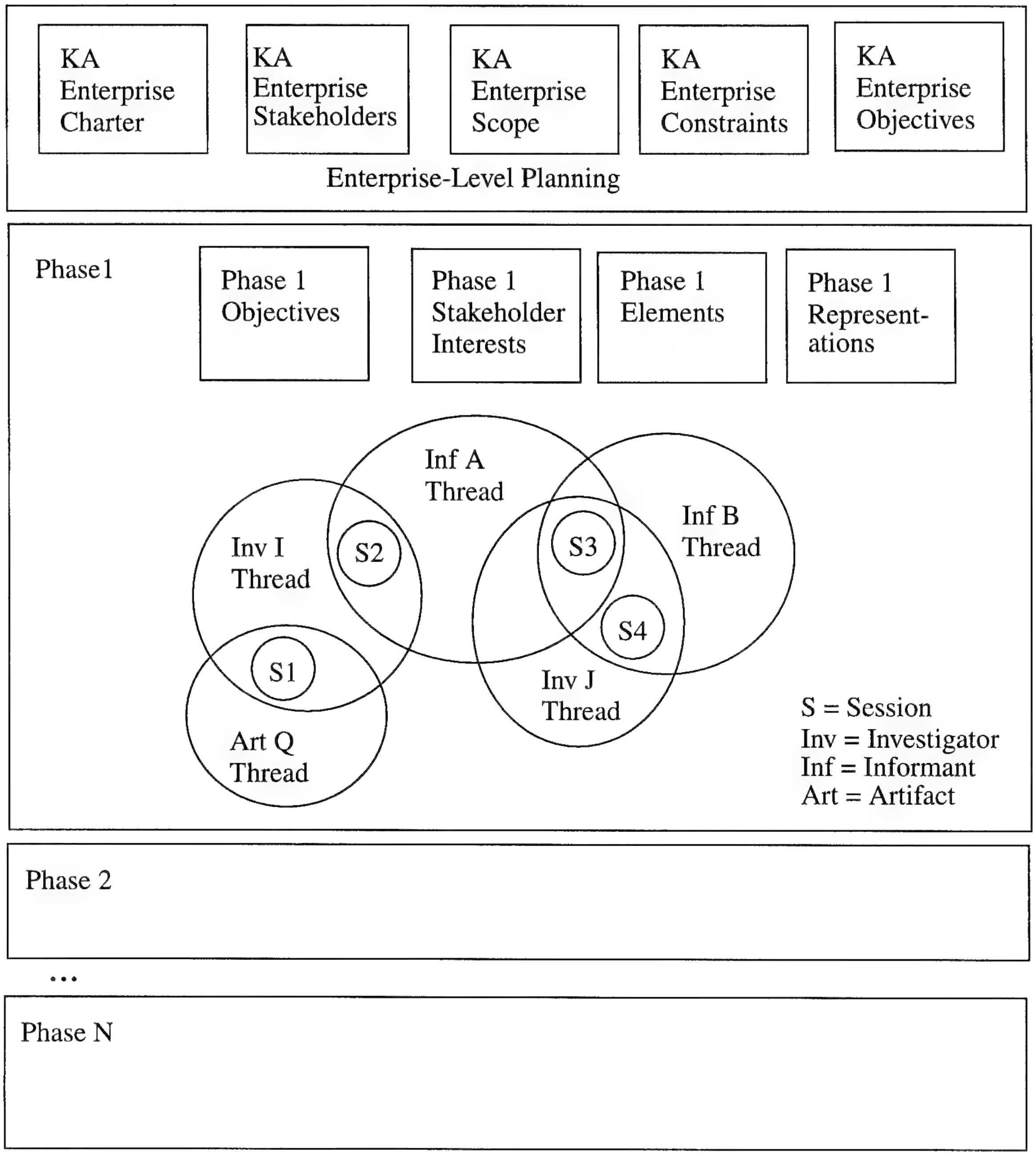


Exhibit 9. KA Plan: Enterprise, Phase, Thread, Session Levels

Enterprise-level planning elements provide a framework for all KA activities and clarify the interface between these activities and the overall project context. The phase planning helps structure the overall KA effort into manageable segments, each with a clear hand-off to an intended audience within the target community. Within each phase, the thread plans weave together the various KA elements (investigators, informants, artifacts and workproducts, etc.) to make best use of resources, optimize informants' time, and keep results coordinated.

4.1 Enterprise-Level Planning

The most important initial purpose of developing a KA plan is to scope and clarify objectives for knowledge acquisition. The KA enterprise is the project-level scope of planning that integrates the knowledge acquisition process into a larger context such as a technology development or domain engineering project. KA activities coordinated as a single enterprise may address many topics, to be explored in many work settings. The enterprise may require a series of knowledge acquisition sessions to cover a given setting with great thoroughness or may rely on a representative sampling of data. As described in Section 3.3.3, if enterprise goals include systematic treatment of variability then a single topic may be investigated across multiple focus community settings.

Planning the KA effort as a coordinated enterprise allows for both separate and coordinated management of KA decisions and resources. The many inter-dependencies among KA activities can be anticipated and managed: e.g., multiple interviews to be held with a single informant can be coordinated; an interviewer may want access to prior interview results in order to plan his or her session. A more systematic approach to planning for these inter-dependencies facilitates scalability of the process to allow for multiple investigators, informants, etc. over a period of time. Detailed planning of threads and sessions will help to ensure reasonable coverage of information acquired for the resources expended, and to avoid the confusion of having multiple unarticulated goals for a single session.

At the same time, KA activities can take on a level of independence from broader project activities that allows for more flexibility. The KA plan should be centrally accessible to all people involved in the KA enterprise (with appropriate visibility provided for different classes of users). Results can be gathered and collated before transfer to the target community. Where the KA effort can align with the stakeholder interests of people in the focus community, KA results can create immediate value rather than being viewed only as a means to the ends represented by the target community's agenda.

4.2 Determining Enterprise Context

The planning approach described here assumes a scenario where an effort called the ***knowledge acquisition enterprise*** or ***KA enterprise*** is initiated within a larger project context, such as a technology development project. The first step in KA planning is explicitly documenting this broader context and its relation to the anticipated KA enterprise.

We first clarify how we are using the term "KA enterprise" in more detail, and outline the types of scenarios where the concept applies most usefully. Subsequent sub-sections discuss the main context elements that need to be described: the enterprise charter, strategic stakeholders, scope and constraints.

A Note on the Term "Enterprise"

We have chosen the term "KA enterprise" to describe the overall scope of a Canvas-based KA planning framework for several reasons. We want to distinguish a KA effort from the broader ***project context*** in which it takes place, to avoid the confusion, for example, of referring to a KA "project" within a software development "project". Most often, KA efforts will *not* be initiated as a separate project, under separate management or funded by a separate contractual vehicle.

At the same time, the term is intended to suggest more structure and coordination than a set of isolated KA activities within a project related by common dynamics and performance aspects; e.g., if a requirements analyst decides to spend some time observing users of the current system. This is a drawback of the alternative term "KA effort", used here on occasion to discuss the KA activities within a project without emphasis on the formal planning context.

In addition, the term "enterprise" suggests those aspects of KA that remain relatively constant over the course of the broader project context. The initial planning steps outlined here aid in identifying and stabilizing these strategic aspects. For example, the intended audience for individual workproducts created from KA sessions will vary; but a target community for the KA effort as a whole needs to be clearly identified at the outset to guide subsequent more specific planning choices. Thus, a session plan will record the intended audience for the workproducts to be produced in the context of an enterprise plan which identifies the target community.

Enterprises vs. Organizations

The term "enterprise" might suggest to some readers a dedicated organizational structure for KA (i.e., as used in the sense of "enterprise solutions" and "enterprise-wide models"). Our use of the term here does not assume, but can encompass, such a structure. The typical structure assumed by the framework is that of an investigator team formed specifically for the KA enterprise within a project, where the investigators may originally be drawn from members of the focus or target communities respectively, or from a specialized community with competencies in KA. Of course, this scenario can lead to more persistent organizational forms. Over the course of the project, the investigator team may increasingly form a distinct community, both of interest (as strategic stakeholders) and of knowledge (as an emergent community of practice). This may create interest in continuing the KA role beyond the scope of the initiating project. Thus, the investigator team will most typically begin as a strategic stakeholder grouping, and may over time become a distinct community of practice.

By contrast, note that typically the focus and target communities are *not* formed specifically for the KA enterprise. Knowledge acquisition in the Canvas framework assumes pre-existing focus as well as target communities of practice. However, a subset of the focus community (the informants) and of the target community (audience) will be selected as part of the planning process and in the final dissemination of KA results. These sub-groups may gradually take on their own perspectives as strategic stakeholders, changing the roles of the participants within their respective communities. In addition, as with the investigator team, these sub-groups may gradually take on aspects of distinct communities of practice.

Recognition of these dynamics can help planners to anticipate a number of issues in managing the overall project, and to facilitate development of the various emerging communities of practice (investigator, informant, audience) in ways appropriate to overall goals.

In some cases an ongoing community of practice functions in an investigator role for repeated engagements. For example, groups have been formed within large companies such as DEC or Xerox to provide specialized KA services as members of design teams for development of spe-

cific technologies. In such cases one could speak of a "KA enterprise" in a more literal sense, as a KA organization. Here, investigators have a stake in developing their capabilities for repeated KA engagements, even in domains where they do not have specific expertise. Addressing the full strategic requirements of such a KA organization is beyond the scope of this document, although Canvas concepts would be of value in determining the mission and scope of such an organization.

4.2.1 Defining KA Enterprise Charter

The ***KA enterprise charter*** documents the project context within which KA will take place. The following questions provide a checklist of issues to consider:

1) What is the broader project context within which KA will take place? What type of project is it? Is the concept of systematic knowledge acquisition an accepted aspect of the project, or does it represent an innovative approach that will require some explanation and education?

2) Where will KA activities take place within the broader project context? Will all these activities be coordinated within the framework of the KA enterprise?

3) What are the key inputs available from the broader project context to the KA planning (and later, performance) activities?

4) What are the results expected from the KA enterprise? How critical are they to overall project objectives? Are there additional "spin-off" results expected from the KA enterprise that do not directly address broader project goals?

One helpful way of documenting the context is an overall process diagram, organizational picture, and/or life cycle model for the project, with indications of where the KA effort falls within that picture. In Appendix A, which describes the interface between ODM and Canvas, Exhibit 26 provides an example of this kind of process diagram, showing the place of the knowledge acquisition enterprise within the ODM domain engineering life cycle.

Different Project Contexts

For the focus of this document, the most typical scenario would be a technology development project for which planners have decided to approach certain information-gathering activities from a coordinated KA perspective. This scenario would correspond to KA's place within a broader SEP or ODM context, for example.

> Example. Viewing TCIMS as a KA enterprise, the overall project scope was a large-scale consortium program involving multiple companies, potential user groups, etc. to develop next-generation medical technology, covering a range of potential technologies including conventional software systems and intelligent systems. Potential applications included trauma care situations in military (both combat and humanitarian or peace-keeping missions), and civilian (urban and rural) settings.
>
> Knowledge acquisition aspects of the overall program were structured as a separately managed project (probably to an atypical degree). Separate organizations and individuals were tapped as investigators. These individuals had varying degrees of specialized knowledge in KA and knowledge representation/modeling techniques, and varying degrees of knowledge in the medical fields of interest. They thus had varying degrees in continuing a formal KA role after the completion of the TCIMS project itself.

KA activities could play a role at multiple points in the life cycle of such a technology development project. A requirements analyst developing specifications for the new system might interview end-users about particular requirements or problems with the current system. For deployment of technology in environments where some legacy technology is already in place, KA can be used as a way of understanding the current workplace and the role of technology in that workplace.

In each of these cases, failure to utilize a KA perspective results in increased risk and lost opportunities. Gathering information about people's desires, expectations, or guiding decisions may introduce very different dynamics as compared to acquiring knowledge about current practice. Activities viewed by developers as "requirements gathering" may be perceived by participants as knowledge acquisition, thus raising numerous stakeholder issues and points of resistance (e.g., protection of knowledge to preserve job security). Or real KA activities may be interpreted as requirements gathering, and desires expressed in knowledge-gathering sessions may take on the problematic status of requirements requested and not met by developers.

The above points concern transfer of information *about* the operational environment where the system will be deployed *to* system developers. Some knowledge transfer is also required *from* developer *to* the user community, in terms of installation support, training, technical support, and acceptance testing. The extent to which the deployed system is a "cultural artifact" of the developers' community, a community distinct from that which will make use of the system, is often nearly invisible to the developers themselves.

Use of a KA framework as a way of gathering "as-is" information prior to intervention is not confined to technology development in the narrow sense. Similar information-gathering or assessment could be done by a consultant as preparation for recommending changes in business practices or work processes.

A common theme in these scenarios is that the KA enterprise provides information to guide intervention in the focus community. In certain cases no obvious intervention, such as deployment of a new system or changes to work processes, is intended for the focus community. In such cases the "broader project context" might appear to be almost coincident with the KA enterprise itself.

> Example. Consider knowledge acquisition undertaken as part of a student's graduate research project in the social sciences. Such a project is motivated (and financed) as part of a larger theoretical agenda, probably involving multiple researchers over a long period of time. Here the objective of knowledge acquisition is to codify some knowledge from a particular community and make it accessible to a different community, of academic researchers. The research results may also turn out to be of interest to the community being studied; and in any case the wider exposure of that information may dramatically affect that community.

Thus the same issues are involved, though they may be veiled beneath the subtle political and professional "stakeholder" interests in academic research. Even in such cases, where intervention in the focus community is not an explicit goal, it is still useful to consider the ultimate stakeholders for the knowledge acquisition effort.

4.2.2 Identifying KA Enterprise Strategic Stakeholders

The KA enterprise charter helps to establish the primary ***stakeholders*** for the enterprise. The purpose of this activity is to identify all groups or individuals that have particular strategic interests in the process or outcomes of the KA activities to be performed. The eventual objectives determined for the KA enterprise need to be aligned with the interests of these key stakeholders. In the following paragraphs, we clarify distinctive aspects of stakeholders for the KA enterprise.

Who Are Stakeholders?

While the notion of stakeholders has general relevance for any project planning, we will need to consider stakeholder roles and relationships from a few aspects particular to the KA context. In general, stakeholders can include the following types of entities:

- Organizations
- Projects
- Groups (ongoing, or chartered with a specific mission or task relevant to the overall project)
- Individuals (project managers, key decision-makers, potential project members)

The key characteristic of entities described in the stakeholder view is that they have *strategic* significance. Later we will map the organizations and individuals involved in terms of their knowledge in various areas. Here we are looking for strategic relationships such as sponsorship, accountability, authority, potential competition, and provider-customer relationships.

The stakeholder roles that are most significant from the standpoint of knowledge acquisition are those that have already been discussed in some detail:

- *Who <u>has</u> the knowledge* that is the focus of the KA effort? These people will be candidates for becoming informants as part of the KA enterprise. However, those people who are selected as informants are not the only stakeholders from a strategic perspective. Anyone in the community who self-identifies or is identified as a "holder" of the knowledge is likely to be affected in some way by the outcome of the project.
- *Who* <u>wants</u> *the knowledge?* These people will be candidates for the audience for the codified knowledge within the target community. Once again, those who are not selected for the audience will also have a stakeholder perspective to be considered.
- *Who will <u>elicit</u> the knowledge?* These people are the candidates for becoming investigators for the KA enterprise. They may have many stakeholder issues of their own.

This step may cause some confusion if we imagine carrying out the steps outlined here in a strictly sequential way, as it seems to be based on decisions not yet made: i.e., who are the specific informants, audience, investigators, etc.? The primary purpose for doing this analysis at this point is to translate the stakeholder perspective of the broader project into KA-specific terms that can inform the rest of the planning process. After specific KA roles have been selected for particular communities, further rounds of stakeholder analysis will be performed as necessary.

In addition to identifying KA-specific stakeholder roles, it is important to scan here for *multiple* stakeholders that could potentially perform the same role. For example, many potential groups could be a source of informants. This may necessitate prioritizing and selection later on; or, if gathering systematic variability information is an explicit objective, it may be necessary to gather data from multiple groups in a manner that facilitates comparison and synthesis of this data.

Depending on the type of overall project context and the degree of formality with which it is performed, there may be more or less thorough stakeholder information available to draw on to determine the strategic stakeholders for the KA enterprise. For a software or system engineering effort, a formal stakeholder analysis is not a typical up-front planning activity (although it is a good general risk-reduction strategy). Thus there may be some "remedial" stakeholder analysis required. In a formal ODM context, a model of domain engineering project stakeholders would be created as part of the initial Domain Planning phase.

The emphasis here is on tracing the broader project stakeholder picture down to the stakeholder aspects relevant to knowledge acquisition. However, this does not mean that the KA enterprise stakeholders will always be a subset of the broader project stakeholders (even one further differentiated in terms of KA roles). To the extent that the KA enterprise charter has a general knowledge area as a goal, KA stakeholders may include stakeholders not considered at the broader project level. In effect, a topic or knowledge area creates a focus for a community of practice that will span organizational boundaries.

> Example. Consider a project to transfer technical expertise from a research and development (R&D) group into a codified form that can be utilized by strategic marketers and proposal analysts to help them identify candidate research projects to pursue. Here, the overall project context could be a specific response to identified proposal opportunities, or a more general initiative to institute more effective knowledge management across groups (hence, an internal organizational change or process improvement project). However, any topic selected for a specific KA effort will probably involve a much broader research, academic and industrial community, of which the research staff available to the project are only a small subset. From the knowledge acquisition planning aspect, this broader set of stakeholders must be considered because the entire strategic value of the KA effort may rest on an understanding of the internal researchers' position within this broader community.

This helps to highlight a distinction that is important to the KA planning process: between *stakeholder groups* and *communities of practice*. The former have common strategic interests; the latter have a common basis of knowledge about some area of practice. The two ways of viewing people and organizations are complementary but distinct. Not every stakeholder group will form a distinct community of practice; the group may be drawn from a larger community of practice, or could link people from multiple communities. Conversely, for a given KA effort, not every community of practice will have a relevant stakeholder interest. This distinction will be brought out at various points further in the remaining planning process.

Documenting the Information

Generally this type of stakeholder information is familiar to the point of being "obvious" to many project members. Since much of the information is "ready to hand" a simple brainstorming approach can be effective, preferably with a small group. At this point inclusiveness is a good criterion, so iteration should result in additional stakeholders missed on previous rounds. Thus flipcharts are a reasonable medium, because the information, once captured, usually stays around (not a lot of deleting) and the chart is easily archived.

Stakeholders listed may be clustered according to organization or project-specific roles but it is difficult to simultaneously gather the names and develop a semantically significant spatial placement of the data. If deemed useful, a second pass over the listed stakeholders can be used to transcribe the material into a more clustered or spatially organized form.

4.2.3 Determining KA Enterprise Scope

The KA enterprise scope is a further refinement of the charter and strategic stakeholder picture for the enterprise. If the charter provides the organizational context, the enterprise scope provides the *thematic* or *topical* context for the KA enterprise. The scope helps to determine what overall topics or domains of interest are appropriate for the project. The charter may identify overall organizational objectives or goals, but these may be satisfied by a variety of topics as focal points for the effort. Typically, though, there are topics that have been pre-determined to some degree as the intended focus of interest. The purpose of this step is to make sure that these assumptions have been explicitly documented.

Example. In the TCIMS project, it was clearly an objective to capture knowledge useful to technology developers in creating new applications to support trauma care. However, knowledge about this topic is closely interwoven with knowledge about many other areas of the medical field, many interacting systems in the operational environment, etc. This meant that some data was collected that was deemed not useful by developers.

Determining the scope may appear to be a "nicety" in the planning process. Based on initial case studies and our subsequent experience, we believe failure to clarify scope can lead to significant breakdowns in the overall project downstream. The reason is that the KA process becomes the pressure point for conflicting stakeholder objectives.

For example, in a conventional software engineering project, KA can be used as a kind of "pre-requirements" data gathering phase. But what, then, is the scope of the knowledge to be gathered about the current work environment?

- Workflow information: how people perform tasks, collaborate, make decisions?
- System interaction information: how people interact with the system to get their job done?
- Desired improvements, changes, complaints: all the detailed knowledge about how systems *don't* work and how people manage to work around them?

Gathering too much of the wrong kind of information for the developers' needs will consume scarce resources, reduce interest in KA results by increasing the noise-to-data ratio, and risk creating inappropriate expectations on the part of people in the user community. Yet without a definite set of requirements, the KA investigator is caught in a Catch-22, because there may be no obvious way to translate the developers' interests into a clear scope for managing the data gathering process. Lack of clarity around fundamental questions—whether "as is" or "to be" information is to be gathered, system or workflow, routine process vs. intelligent decision-making, individual vs. group activity, whether exhaustive or sampled data gathering is required—can create major problems.

In the ODM context, knowledge acquisition does not begin until there is a clear and validated domain definition established. Recognition of the issues above is one motivation for this process structure within ODM. Because domain engineering involves the comparative study of multiple systems or potential settings for new systems, any problems in scope that beset conventional single-system engineering are likely to be magnified in the domain engineering context. ODM places KA within a clearly defined *descriptive* phase; this structure helps communicate to informants that gathering information is a different process from acting upon the information as requirements.

This same descriptive framework can be useful for any KA effort performed as part of a larger technology development project. In fact, the ODM processes of domain scoping and definition, as described in detail in [49], can provide a useful approach in any KA effort.

Example. In the TCIMS effort, certain high-level domain partitions were established to help structure and guide the KA enterprise. These included a focus on military, civilian urban, and civilian rural trauma care situations in separate phases of the project. The focus on trauma care defined a scope that excluded other areas amenable to the core technology approach (hand-held networked computer technology in the field), such as remote monitoring of patients with chronic medical conditions. Such diverse application environments could be investigated by a marketing study but would not provide an adequate focus for a KA effort.

4.2.4 Identifying KA Enterprise Constraints

Constraints are a factor in any project planning effort; but, once again, our focus on knowledge acquisition raises particular constraint issues that should be explicitly addressed. These issues include the following:

- *Constraints on informant availability.* Informants' time is a scarce resource. Unlike typical development projects, the norm for a KA effort is that informants will not be assigned as KA staff; instead they must be "borrowed" for sessions from their own management, active projects, etc.

- *Restrictions on data.* By its nature, knowledge is a sensitive area of intellectual property. Yet the business of KA, as defined here, is transfer of knowledge across communities of practice. There can be stringent restrictions on the free propagation of such knowledge, both outside of and within corporations. Issues such as use of data of classified status, proprietary data in consortium situations, legal ramifications of inaccurate data, or data that raises personal privacy issues (e.g., medical case histories) need to be considered.

- *Budget restrictions.* One of the major advantages of enterprise-level KA planning is that it helps ensure an efficient use of resources for KA activities. This requires some notion of the overall budget of resources within which to search for reasonable strategies. In addition, knowledge acquisition, unlike system design to a specified set of requirements, is by its nature an "open-ended" activity. It is difficult to say when one has gathered "enough" knowledge about a given topic. Identifying constraints on effort (and monitoring effort expended) are one fairly objective way to maintain control of the process.

  Here one of the important points is to consider the "invisible" time sinks in a KA project. As a simple example, for the TCIMS project the typical time required to write up a thorough report of a one-hour interview was on the order of eight hours.

- *Time constraints.* For many KA enterprises, the primary goal is to get certain information into the hands of people in the target community in a timely fashion to help support certain interventions. There may be hard windows of opportunity that constitute successful transfer of data for the enterprise. These must be clearly laid out up front.

  In situations where KA is being used to support technology efforts, it can happen that the schedule pressures on developers force them to begin work on the system long before the results of KA are distilled and ready for hand-off. Just as the investigators often have no direct managerial access to informants, so they must often respond to these other schedule constraints or risk having their work ignored.

Each constraint category outlined above has some particular significance for KA. It is useful to document these constraints as a final step of the KA enterprise context-setting part of planning.

4.3 KA Enterprise Objectives

The preceding section suggests ways of identifying overall project goals that guide the KA process. These goals must be translated into actionable plans for KA: specifics about where and from what sources knowledge is to be acquired, how it is to be represented, etc. This section outlines a process for deriving these detailed objectives for the KA enterprise.

To understand the steps of this part of the KA planning process, it is helpful to characterize what a final "KA objective" will look like. Linking back to our original definition of KA as a transfer of knowledge across communities of practice, we will need to identify the following:

- Which communities of practice are involved in the transfer?
- What kind of knowledge is to be transferred?
- What will be done with the knowledge? That is, what will someone in the target community be able to do, with the transferred knowledge, that s/he could not have done previously? (Without answering this question we can't establish success criteria for the transfer.)

These questions are addressed, first, by identifying the distinct ***communities of practice*** that fall within the strategic scope of the KA enterprise. This process is described in detail in Section 4.3.1. These communities are referenced in the second main component of the objectives, the ***knowledge transfer configurations***, which directly reference the distinct KA "modes" introduced in Section 3.1.6. Each configuration places particular communities in the focus, investigator and target roles respectively, and establishes specific knowledge transfer objectives for those communities. Knowledge transfer configurations are described in detail in Section 4.3.2. There may be multiple objectives defined for a given KA enterprise; these objectives may resolve into different phases of the overall effort. The third component of the KA objectives involves global objectives such as the extent to which systematic acquisition of variability data is required.

4.3.1 Identifying Communities of Practice

At the start of planning we developed a strategic stakeholder picture which helped to establish the charter, scope and constraints for the KA effort. We now need a more detailed view of the various distinct communities of practice within the scope of the KA enterprise, and their relationships from the standpoint of potential knowledge transfer. This planning step builds directly on the core concepts introduced in Section 3.1.1. It provides a basis for defining specific KA objectives in terms of the knowledge transfer roles to be played by these communities.

Use of the term "community of practice" as part of the Canvas approach emphasizes our orientation towards the type of knowledge that is created and maintained through collaborative practice in the work setting. It is important to distinguish strategic stakeholder groups from communities of practice; these can be, but need not be, coincident. In general, communities of practice are *ongoing* communities, not defined by tasks or projects with fixed time frames. This is partly because knowledge and shared culture grow over longer periods of time; partly because a community not defined by a single project goes through different kinds of formation, boundary maintenance and focusing dynamics.

So an organizational entity considered as a single strategic stakeholder for planning purposes may span multiple communities of practice; conversely, a single community may have multiple strategic "stakes" in the KA enterprise. Stakeholder vs. community-of-practice relations are interdependent but separable. One administrative unit of an organization may house several distinct communities of practice, perhaps as a result of prior mergers or organizational restructuring, perhaps as a result of stratification of roles (e.g., engineers vs. marketing personnel). Conversely, a single community of practice may span several groups called out as distinct strategic stakeholders and multiple organizations.

Communities of practice can be conceived as overlapping or even in subset/superset relationships.

> Example. Consider a hospital environment, where doctors and nurses work together with an extensive shared professional culture. Doctors, nurses, and other personnel such as administrative staff can be said to form a distinct community of practice. At the same time, doctors have specialized terminology, social and cultural interactions which make them a distinct community of practice; similarly, nursing staff would form a distinct community.

The distinctions between stakeholder groups and communities of practice are not hard and fast. A group of people chartered to carry out some decision-making task in a very specified and constrained time frame are best dealt with as a strategic stakeholder group. If, however, this group persists in time long enough to begin to form specific group dynamics, such as boundary criteria, shared language, values and objectives, the group may begin to act from a base of shared knowledge.

A transitional notion could be termed a "community of action". Any group with shared knowledge has some shared context for action utilizing that knowledge that helps shapes the dynamics of the community. When the organizing principle of the group is primarily around shared tasks or activities without a common and distinguishing body of knowledge as a foundation, we can consider the group more a community of action than a community of practice.

> Example. In the hospital environment, a group of doctors and nurses are convened as a task force to oversee the adoption of a new computer system. As the task force begins, the group represents a collection of different stakeholder views, linked by a common strategic perspective (they are all members of the task force). All members of the task force may be members of the hospital staff community of practice, but that does not make the task force a *distinct* community of practice. It is, however, a "community of action".

If, over time, such a group develops shared context around the system evaluation process, interactions with system developers, etc., they may increasingly take on a distinct community of practice role. Conversely, once communities of practice become involved in a KA effort, certain unique stakeholder interests arise out of that participation. Thus, later in the KA planning process, we will revisit strategic stakeholder issues again (as described in Section 4.4). At this point, KA enterprise roles will be defined in terms of membership in specific communities of practice. Membership in a community of practice with a defined knowledge transfer role (i.e., focus, investigator, target) will engender specific stakeholder interests that need to be considered in more detail.

Community of Practice Diagrams

For capturing a "communities of practice" view, we describe a specific technique developed in a pilot application of the Canvas planning approach. Readers should feel free to try this approach, experimenting and extending as appropriate.

The basic notation is that of simple Venn diagrams or "interlocking bubble" diagrams. Each bubble or circle is labelled with a name for a specific proposed community of practice. Where possible, choose names that help maintain a separation between strategic or business entities and the groupings that correspond to communities of practice. It is possible to build several layers of the diagrams, to describe "sub-communities" where it is necessary to look at the relations in more detail. It may also be helpful to "populate" the communities with example individuals, represented as "dots" (in a separate color) placed within the interlocking circles as appropriate. Of particular interest are "bridgers", individuals or groups who are placed in the overlap of two communities. As a validation step, it should be possible to name at least one such bridging individual for any communities drawn as overlapping.

This diagrammatic form is particularly amenable to development and validation in collaborative group sessions. Rich data can emerge in the process of developing a communities of practice view and deciding, for example, whether two communities are disjoint or overlapping. The act of recording and discussing the data stimulates re-thinking; thus frequent changes and re-drawing may occur. To facilitate this form of elicitation, the diagrams are best drawn on a medium facilitating common view during creation, lots of space, and the ability to rub out and correct, i.e., a whiteboard rather than a flip-chart for the initial draft.

The activity of capturing the communities view will often elicit historical data. For example, when people from different groups are moved together, a legacy of distinct communities often remains even though a new overall culture will form slowly. Similarly, when geographic dispersion occurs a shared community may persevere for a while, although eventually the separated groups will tend to form different communities. These change processes are slow relative to project life cycle, and would generally occur over a span of at least several projects.

In addition to static relationships like inclusion and overlapping, and the historical data that may underlie these relations, certain patterns of knowledge transfer among communities may emerge from the analysis. These may foreshadow the knowledge transfer configurations that will be determined as specific objectives for the KA effort, but they are more descriptive in nature. For example, one common pattern is a "chain" pattern that corresponds to well-known technology transfer roles (innovator, transfer agent, receptor, adopter).

To summarize, the communities of practice view should indicate the following:

- Relevant communities and sub-communities within the scope of the KA enterprise. This should include focus, investigator and target communities.
- Relationships of subset/superset between communities. In these cases, validate with a specific individual in the broader but not the narrower community.
- Relationships of overlap. Validate by identifying a "bridging" individual or group that is a member of both communities.
- Layering of diagrams where appropriate so that a given sub-community can be described in a separate picture.
- Consideration of historical relations that may underlie the formation or divergence of communities of practice.
- Patterns such as innovator/transfer agent/receptor/adopter chains.

There are other potential uses for the community of practice view of a given organizational context. For example, such a view can be used to document, and possibly to predict, where potential communication breakdowns may result from discontinuity in underlying patterns of shared versus localized knowledge. These types of breakdowns are distinct from those resulting from individual temperament or interpersonal conflict; they are, as it were, consequences of boundaries in the "knowledge environment". This sub-section has focused only on capturing the information about community of practice relations needed to establish KA enterprise objectives.

4.3.2 Identifying Knowledge Transfer Configurations

A ***knowledge transfer configuration*** shows one way that communities of practice can be organized into focus, investigator, and target roles. Some of the possible basic knowledge transfer configurations were shown generically in Section 3.1.6. The purpose of this activity is to identify all potential configurations of this kind within the enterprise scope, working from the enterprise charter, strategic stakeholders and constraints, and utilizing the communities of practice view developed in Section 4.3.1.

This activity can be done in "brainstorming" mode because the immediate goal is not to decide on a specific configuration but to generate a number of alternatives. Each alternative should be relevant to some identified goal for the overall project, but in general more configurations will be identified than can be addressed given the resources available for the KA effort.

It is necessary to identify target and focus communities for the KA effort. If informants are considered to be potential users of the information gathered, then the target community includes the focus community. If there is a third, distinct target community (e.g., technology developers) then the investigators will be playing a strong "translation" role between the communities. For each knowledge transfer configuration identified, a knowledge transfer configuration diagram is drawn.

Knowledge transfer configuration diagrams

Here we introduce another suggested diagramming technique for experimentation, based on some initial trial experimentation. Knowledge transfer configurations can be notated as bubble-and-arrow diagrams with some simple semantics. A bubble represents a community of practice, generally decomposed to the smallest grouping relevant to the objectives. These diagrams are essentially sequential lists of diagrams based on the knowledge transfer modes described in Section 3.1.6 and illustrated in Exhibit 4.

Three types of bubbles need to be distinguished: focus, investigator and target roles respectively. A given bubble is labeled with the name of a distinct community of practice (from the communities of practice diagram) that fills the role. Arrows are drawn from the focus community to the investigator community (representing an elicitation transfer), and from the investigator community to the target community (representing codification and transfer).

In a given configuration, a single community may play multiple roles; in this case, we draw the community bubble only once. For example, a community playing both investigator and target roles is depicted with a "self-referent" arrow from and to the same bubble. Several bubbles can play a given role in a single diagram; for example, when an investigator community integrates knowledge from two or more focus communities.

Nested diagrams can be used to expand bubbles to a lower level of detail, i.e., showing more specific transfer among sub-communities. Configuration diagrams should not, however, be decomposed down to the level of dealing with knowledge transfer to specific individuals. The KA process is more than the learning of a new field by a single individual. Thus enterprise objectives should emphasize, and the knowledge transfer configuration diagrams should illustrate, the enabling of knowledge transfer across communities as a whole, not merely individuals within those communities.

Configurations can also overlap. For example, a community playing the target role in one configuration can later play an investigator role for a new target community farther down the knowledge transfer "pipeline". We do not attempt to depict these temporal relations between configurations at this point; this would overload the simple semantics of the diagrams that make them useful for clarifying objectives.

The configurations are used directly in selecting specific objectives for the KA effort. In addition, the process of developing the set of configurations is useful as a form of risk assessment and to flag relevant stakeholder interests to consider. The insights generated by the brainstorming and revision process are often worth noting for future reference. Refinement, consolidation or splitting of configurations is part of the process. First-draft diagrams will be superseded by latter diagrams, where the scope of various roles (source, investigator, target) may have been widened or narrowed.

Eventually, the transfer modes introduced in Section 3.1.6 could be extended into a set of patterns with associated guidelines and issues to consider. For example, a typical configuration is when the investigator community is a sub-community of the target community; e.g., a group of software engineers are given the task of collecting end-user requirements for a new system. Issues to con-

sider might be: barriers to dissemination of the knowledge acquired to peer groups within the target community that may be identified or self-identify as competitors to the investigator group; bias introduced by the implicit contextual assumptions shared between investigators and audience; and the risks of the investigator community being inadequately trained in the specific skills and practices of knowledge acquisition as a discipline.

4.3.3 Selecting Specific Objectives

Once knowledge transfer configuration diagrams have been created, specific objectives can be selected for the KA effort. First, a set of knowledge transfer configurations are selected as goals for the effort. Knowledge transfer configurations should be selected that meet stakeholder expectations and are consistent with the KA enterprise charter and constraints.

Each objective should have these components:

- Reference to a specific knowledge transfer configuration as the "map" of the knowledge transfer goals;
- Reference to a specific domain, topic or range of topics/types of knowledge that are the focus of the transfer;
- Criteria for the intervention to be effected by the KA process relative to this objective;
- Specific objectives for each stakeholder group and/or community of practice involved in the configuration;
- Traceability of the objective to overall project goals (as reflected in enterprise charter).

4.3.4 Developing Phase Plan

If the KA effort is large enough for knowledge acquisition to be performed in a number of phases, knowledge transfer configurations can be selected for each phase. Further planning activities described in this section may need to be revisited for individual phases, including setting phase objectives, assessing stakeholder interests, selecting elements, and planning threads and sessions. These activities are not qualitatively different at the enterprise or phase level, so the discussions can be applied readily where appropriate.

It is helpful to have clear differentiation between phases. Different topics or domains of focus are one obvious way of breaking the enterprise into phases; but these should be traced, if possible, to differences in the communities of practice involved—different focus communities, different target communities, etc. One typical "roll-out" pattern for phases involves moving the entire KA life cycle "downstream", so that investigators become the "answer people" or surrogate informants, and the audience within the target community become transfer agents to a further community.

When you have considered all configurations, some will be allocated to various phases of the phase plan. Others will be out of scope for the current planning effort, but might be noted in a "future opportunities" list for later consideration. Not all phases need to be committed phases; one of the desired outcomes of a given phase can be to obtain commitment from key stakeholders to proceed to some later phase. However, at the start of a given phase, it is necessary to know whether or not KA will be carried out on the assumption that later phases will go forward. This can have a significant impact on resources allocated and the way KA is actually performed.

As a rule of thumb, each phase should have as an outcome effective transfer of some knowledge to a target community, where this transfer creates value for the target community (i.e., is not merely a means to accomplishing a further KA objective). The primary stakeholders for each phase should be clear. One additional advantage of the phased approach, besides breaking the enterprise into more manageable phases, is to identify as many of these "interim outputs" as possible for the enterprise.

When multiple configurations apply within a given phase, there should be clear priorities for which have primary importance, and which are of secondary or "nice to have" status. Some configurations may capture interim requirements, such as the need to obtain validation from the source community. Such validation may impose requirements for specific workproducts or notations to ensure the ability to get validation, but are not worth treating as a separate project phase in terms of overall planning. They are necessary steps to get the data captured and validated.

Within each phase, there will be a common cycle of activities: planning, orientation, detailed data gathering, integration, codification, presentation to the target audience, and re-planning. Several of these cycles may occur within a given phase.

Planning a Pilot Phase

In our pilot application of the Canvas planning approach, resource constraints permitted only a very cursory execution of the KA data gathering activity itself. Our experience suggests that there are general advantages to this approach as a first phase of knowledge acquisition, even (or perhaps especially) for large KA enterprises. A short pilot of one to three weeks' duration serves to validate the basic outline of the plan, get the investigators some direct experience in the field if this has not already happened prior to formal planning, and provides a shake-down of the dossier structure. Audience requirements gathering activities can be integrated into the initial phase. The latter point is of particular importance; if the dossier structure is ill-formed, then the more data that is collected before adjusting it, the greater the risk of potential loss of data or wasted effort due to rework.

The pilot can also serve as a "calibration" pass to better hone estimates of things like required preparation time for interviews, time required to debrief after sessions, etc. The flexible, responsive "adaptive planning" approach described here, where results of interim KA sessions can reset priorities and plans for subsequent knowledge acquisition activities, requires a high degree of process maturity and estimation skill. Of course, such metrics will improve incrementally over the life of the project, but it can help to identify major discrepancies prior to doing the final setting of expectations for the project with sponsors and customers. Last but not least, by attempting to carry out the pilot in a minimal amount of time, the project is simultaneously minimizing risk and stress-testing the degree of "finesse" supported by the KA project management infrastructure. The pilot should reflect a desired shortest increment for a cycle prior to adaptive re-planning. For large projects, a pilot might even be desirable before seeking sign-off on the overall KA plan.

4.4 Assessing Stakeholder Interests

Once specific objectives have been determined for the KA enterprise and a preliminary phase plan is established, it is advisable to do a more detailed level of stakeholder interests assessment. The knowledge acquisition process has established roles using the knowledge transfer configurations described in Section 4.3.2. Section 4.5 will describe criteria for selecting and characterizing these roles in terms of their potential value to the KA results. Here we consider these roles from a stakeholder point of view.

This second phase of stakeholder analysis serves to validates the objectives established by systematically considering the impact of the desired knowledge transfer results from the standpoint of all stakeholders involved. For each stakeholder, there may be conflicting incentives and disincentives. The planning process should involve building the complete strategic picture before attempting to develop piecemeal responses to particular issues and concerns. We are concerned with the following questions:

- What motivators and/or incentives do stakeholders have to participate in the KA enterprise?
- What barriers, perceived risks or disincentives might exist for their participation? Having identified key stakeholders with sufficient motivation to initiate the effort, it is easy to forget to consider other players that may have strong interests threatened or compromised by the KA objectives. By identifying potential obstacles such as certain stakeholders' resistance to enterprise objectives early in the process, mitigating strategies can be developed.
- What additional or synergistic outcomes could be defined for the KA process that might create added value from the perspective of one or more stakeholders in the process? These can help establish secondary objectives supporting the primary enterprise/phase objectives.

In addition, stakeholder assessment suggests criteria for selecting the specific elements (e.g., individual investigators, informants, artifacts, etc.) for phases, threads and sessions based on stakeholder considerations.

In the following sub-sections we consider the incentives and disincentives relevant to knowledge acquisition for each role in the KA enterprise (focus, investigator and target community) as stakeholders. Examples are cited from experiences in applying SEP on the TCIMS project. In the discussions below, we assume that the *investigators* are those playing the role of gathering stakeholder input. In fact, this role could be played by a variety of people involved with project planning.

The assessment process should include active involvement with stakeholders themselves. This assessment process may thus reflect some elements of knowledge acquisition in its own right. For example, it may require in-depth listening to stakeholders' concerns about project impact or experience from previous projects similar (or perceived as similar) to the current KA effort. Planners must be able to explain how the project differs from other previous research or change projects. This process also involves active communication about the nature of the KA effort and the larger project context to stakeholders. The way in which this assessment and presentation is performed may have a strong influence on the acceptance of, and ultimate success of, the project.

4.4.1 Focus Community Interests

The most salient fact to consider in terms of focus community stakeholder interests is that these are the people who are the source of knowledge for the KA effort. Whether the overall objective is to codify that knowledge more effectively within the community, or to transfer it to members of other communities, the net result is a *change* to those relationships within this community that are defined by relative levels of knowledge.

The impact of this kind of organization or cultural change is easy to underestimate, because in many organizations the most visible structural relations are based on authority, power and accountability rather than knowledge per se. (The fact that Joe down the hall is the person to consult about particular types of problems rarely shows up on organization charts.) Therefore significant changes in the "knowledge ecology" of a group or organization may not appear as organizational change; yet the repercussions may be as or more dramatic. In assessing the stake-

holder interests of the focus community, therefore, we must consider the many reasons why a change in the structure of knowledge relationships may serve, or may threaten, the strategic interests of the various people involved.

Motivators/Incentives

The KA process depends upon available artifacts and a pool of informants. Since informants must take time out of their work to participate in the KA process, it is important to consider what incentives are available to ensure that informants are engaged and enthusiastic about participating in the KA process.

There can be many motivators for informants, including the following:

- potential enhancement of their status in the community as a result of participation;
- participating in a project that will make a useful contribution to their field;
- influencing key players they have been unable to influence;
- being heard, seen and recognized for their expertise;
- opportunity to improve the work environment through their input;
- opportunity to learn from and exchange knowledge with other practitioners;
- opportunity to disseminate their knowledge and experience more widely.

Informants' perceptions of the project will have a major impact on their enthusiasm and willingness to participate. A key incentive is belief that results of the KA effort, and the larger project of which it is part, will make a useful contribution to their community. This depends upon the investigator's ability to positively impress potential informants about the value of the project in question, and the credibility of its anticipated results within the focus community. Informants will want to know that the project is realistic (as opposed to a research "toy"), and feasible (instead of being so ambitious that success is unlikely).

In some situations the potential informants of interest may be in a disempowered position. Thus they may have been trying to influence colleagues, management, policies, etc. for some time and may be exhausted or have even given up. An astute informant may recognize the KA effort as a new and perhaps better vehicle for influencing practice, policy, structure, training, resource allocation, etc. While this can provide strong motivation for participation it is an issue that must be treated with great caution. Investigators may find tension resulting from multiple agendas between project sponsors and informants; e.g., if the KA effort is the most direct avenue for communicating complaints and concerns to management. Investigators may even struggle with their own sympathies since their position is a natural intermediary role. The investigator must negotiate a balance between being an appropriate channel for being heard and being treated as a messenger to ensure benefit to both the informant and to the KA effort.

Similar to the above motivator, a KA effort can provide a forum for listening, understanding, and appreciating employees' knowledge and expertise, and for documenting local, informal innovations. In many organizations, management holds a simplistic view of lower level jobs as being rote and mechanical. Managing by this simplistic view is often less feasible than management would like. John Seely Brown, head of research at Xerox PARC, claims, regarding the informal daily innovation activities of even office clerks, "These informal activities remain mostly invisible since they do not fall within the normal, specified procedures that employees are expected to follow or managers expect to see." [5] A KA effort may be an opportunity for employees to talk to

somebody who cares about these "invisible" activities. The result may be to codify the activities as "knowledge" that will not only improve practice, but will better educate managers in new ways of observing and understanding employees' work.

In general, the KA planner should consider ways of creating incentives for informant participation, including exploring synergistic goals that are peripheral to the main objectives of the project. Even if the KA materials to be gathered are primarily intended for use by a separate community, the KA enterprise may create enough interest on the part of informants that they see uses for the resulting dossier not anticipated in advance (e.g., for training of new workers). The planning process should be flexible enough to explore ways to offer such spin-off value or secondary benefits to the focus community where possible.

Barriers/Disincentives

A number of factors may tend to discourage participation of potential informants, including simple though aggravating factors such as scheduling conflicts or total unavailability. In TCIMS, some informants were reassigned to the field in Somalia and Haiti just as their input was needed, necessitating a restructuring of the informant pool. Expert informants are often in high demand, so scheduling may be non-trivial. Part of the rationale for careful management of dossier materials is to ensure that such experts' time is used as effectively and conservatively as possible.

In addition to these pragmatic issues, there are a number of other real or perceived risks that may act as barriers to informant participation. Informants will naturally be concerned about potential consequences of their participation in a KA effort. There may be distrust of the overall objectives of the project for which KA is being performed (e.g., as part of a "technology push" effort, or part of an undesired trend towards centralization and standardization of practice).

There is a perceived risk in documenting *de facto* rather than "official" procedures. In most organizations, undocumented processes evolve to compensate for weaknesses in formal business structures and approved processes. Practitioners may believe that documenting these processes will cause problems with authority figures in the organization or will invite the scorn of professional peers. Thus they may resist participating in the KA enterprise; or, if they do participate, may "fudge" or dissemble rather than describe the true situation as they see it. In the TCIMS project, there were many such considerations due to both the medical and military context, such as legal issues, military protocol, etc. On the other hand, simply documenting official rather than actual procedures may be viewed as supporting the status quo or evidence of superficiality in data gathering

When such dynamics are at work, it is not enough to use elicitation techniques in the KA sessions themselves as a cross-check for bias. The overall stakeholder picture must be addressed so that informants are not "tricked" into revealing data that may cause problems for them. Strategies to mitigate these concerns must maintain the confidence of the informants.

A related disincentive is the potential of repercussions to individual informants based on information gathered. Responses could include the following:

- Knowledge source information can be kept privileged in specified respects;
- Anonymous aggregation of certain data can be performed. (This is more difficult with the qualitative data typical of KA than with sample quantitative data typical of statistical studies or surveys.)
- KA workproducts can be sanitized (though this risks introduction of systematic distortion in the data).

- The data itself can also be protected as proprietary. For example, In TCIMS, where informants may have perceived such risks for themselves, the consortium agreed at the outset to keep all raw KA data as proprietary, thus guarding to some extent against repercussions to the informants.
- The LIBRA document [48] involving "indirectness" assessment strategies such as gathering data via hypothetical storyboarding versus direct case studies. These techniques can be used to insulate informants from feared political consequences.

Another potential disincentive for informants is the perception that documentation of their knowledge could compromise the job security or advantage offered by their own private expertise. This is certainly a factor in some settings, e.g., the perceived threat that letter sorters might have felt in interviews prior to development of automated letter sorting equipment in the U.S. Post Service. In TCIMS, this did not seem to be an issue for medics, largely due to the knowledge-intensive nature of their specialty. Medics have a common core of practice with an acknowledged high variability in how those core procedures are applied and little likelihood of losing jobs to automation.

The issue above deals with expert informants who want to retain the advantage of their expertise within the community. The KA process can also create tension for informants who lack confidence about their own professional standing within the community. They may feel that participation in the KA process will "put them on the spot" or test their competence in the field.

This potential barrier is one reason we avoid use of the term "domain expert" as a blanket term for a human knowledge source. For many kinds of KA tasks, we want to gather a rich set of data from different perspectives, including those of practitioners in the work setting who would not consider themselves experts in the sense relevant within that community. This issue is particularly important when KA focuses on legacy systems being used in the field. Because of the nature of professional status in technology-intensive cultures, people who interact most extensively as users of computer systems are often assigned relatively low professional status within their work settings. Yet their expertise in dealing with current systems, getting information in and out of them, working around their limits, etc., may be highly relevant to the KA objectives for a technology development project. Thus the very language used in presenting the objectives and structure of the KA effort can build or compromise receptivity.

Other strategies for mitigating this perceived risk, important throughout the KA effort, include: giving informants adequate time to prepare for interviews; making expectations and objectives of sessions clear up front; allowing informants to defer to established sources within the interview for validation ("Check up in the manual on what I said; it's something like that anyway..."); keeping data confidential; and allowing informants to validate the write-up from their sessions to their satisfaction. There are, of course, associated issues with each of these strategies. At a minimum, KA planners must remain sensitive to the double-edged sword that the issue of "expert status" may present to their potential informants.

The above points highlight the fact that stakeholder issues must be considered, not only in planning, but throughout the various interactions with informants over the course of the project. Treating the informant with respect (e.g., being on time, arranging a pleasant environment for the interview, and being personable) may seem an obvious guideline, but can have a great impact on an informant's receptivity to the KA process. Being chosen to participate in a KA effort may be felt to be of questionable value (as may be reflected in humorous characterizations of "participant as victim" or "getting our brains picked"). A KA process that considers informants' needs and interests and values their time and knowledge will go a long way to building receptivity, and even encouraging other practitioners to participate.

Other Focus Community Stakeholders

While the discussion above mostly addresses the interests of those who will be informants for the KA effort, it is also important to remember that all members of the focus community, not just informants, are stakeholders. Some people will resist efforts to codify knowledge, for a variety of reasons. Others may object to the specific target community involved in the effort. In TCIMS, the patients involved in the work settings studied were clearly important stakeholders whose interests needed to be protected, although they were not involved as informants. Patient's rights to privacy in data was a major concern.

The conflicts surrounding these varied interests may stay beneath the surface in the early planning stages, only surfacing around detailed planning issues like the selection of informants. For example, if the informants will get the opportunity to select the knowledge, opinions and viewpoints that are reflected in the KA results, there may be strong political factors involved in who is or is not selected.

4.4.2 Investigator Community Interests

Investigators have their own set of factors motivating their participation in KA that need to be addressed: the organizational structure of the project team, project realities, and individual motivational factors. From a stakeholder perspective, unique characteristics of the investigator role includes the bridging or intermediary function performed, insider/outsider ambiguity in the relationship with the focus community, and the degree of commitment to a distinct community of practice focused on the knowledge acquisition role itself. Investigators are most often drawn from the focus or from the target community and rely on their own relevant knowledge in the field as part of their basis for competency. In some cases there may be an external, professional KA community involved. The stakeholder issues will be different in each of these cases.

- For informants-turned-investigators, a change of status may result which could be viewed as a benefit (a chance to change job roles or be promoted, to gain new skills, to gain visibility) or a risk (getting pulled off of higher-profile active work tasks, losing rapport with one's work community).

- Conversely, for target community "draftees" to the investigator role, there may be mixed incentives and concerns. For example, in a technology-oriented organization involvement in the data-gathering activity may be deemed of lower status, even signalled by different job titles and job descriptions; e.g., are investigators "engineers", or paid as such? For other people such involvement could represent a technical stretch; for example, technical documentation groups may be excellent resources to tap for investigator staff.

- For KA specialists (e.g., researchers, consultants, professional knowledge engineers, ethnographers) a very different set of stakes arises. They may be concerned about maintaining "ownership" of the KA process and method expertise on the project, or may be oriented towards the effective transfer of those skills to other, internal investigators as an objective of the project. They may also have a separate stakeholder agenda of improving their own practice or deriving case studies for research. Developments that pull them too deeply into the domain knowledge of the focus community, e.g., as "surrogate informants" to technologists may present risks to their overall motivation for participation.

We will continue to draw upon TCIMS experience in the medical domain to provide some examples of these issues.

Motivators/Incentives

In TCIMS, investigators were referred to as "KAs" or knowledge engineers. Because of their unique relationship with the informants (called "experts" in the TCIMS context) who were in many cases also decision makers, the KAs became some of the most effective ambassadors for the TCIMS project to the focus community (sometimes called the "stakeholder community" in the TCIMS context). This benefit of being on the front lines with the stakeholders was a source of gratification to some KAs.

Also significant to both KAs and informants was the incentive to see the KA process provide tangible direct benefits to the focus community. In cases where the KA team is serving as intermediary between domain practitioners (e.g., medical personnel) and technologists there is probably a natural tendency for KAs to adopt a bit of a "champion" role, to ensure that the interests of the focus community are being adequately addressed in technology development.

A source of both frustration and gratification for the investigators in TCIMS was that they were drawn more into the technologists' world than they desired. Some interactions between KA personnel and technologists amounted to their review of technologists' designs with respect to informants' expectations as the investigators understood them. This activity did enhance respect for the investigators in the technologists' community, and resulted in an emerging role for the KAs as "ersatz users" or "answer persons" for the technologists, people who could provide information about what would really happen in the field, could partially make the bridge to the technologists' terminology and were more readily accessible than experts in the field.

Barriers/Disincentives

The time and effort required to serve in the role of the "answer person" may also turn out to be a disincentive to some people. The investigator that prefers contact with the focus community may find contact with the technologists a distraction. Also, if a goal of the investigator or the organization the investigator serves is the development of transferable KA expertise then becoming the "answer person" may not be desirable.

The above motivating factors depend upon the investigator achieving a reasonable level of knowledge and understanding about the focus domain. The notion of the investigator thread in the Canvas framework explicitly models the additional domain knowledge acquired by investigators in successive KA sessions, with the potentially both beneficial (e.g., familiarization) and undesired (e.g., bias-producing) effects of this knowledge.

There are other disincentives to the investigator role. The practice of knowledge acquisition is not a commonly accepted role in current technology projects, save those heavily influenced by performer experience, such as in artificial intelligence or similar technology. The result is that investigator is considered "optional" in the hierarchy of project personnel, a major disincentive to becoming an investigator.

Another significant complication for TCIMS investigators was related to project pragmatics and schedule. Investigators want their results used. The most tangible evidence is handoff and use of KA results "downstream" in the project (by technologists). In theory, KA results should have provided input to the requirements specification phase of technology development. Given the iterative nature of the TCIMS project approach, the investigators on TCIMS were of necessity working *in parallel with* the technologists. These technologists were largely comprised of designers and developers of computer software and mobile computing systems. The parallelism of work meant that the investigators were continually adjusting their agendas to meet needs and expectations from both focus and target communities.

This placed considerable time constraints on the TCIMS investigators, pressing them to create early results that could aid the requirements phase of the development effort. This created difficulties, especially considering that the effort needed to adequately document sessions is considerable, being several hours per actual "contact hour." Schedules were further strained by the difficulty in getting access to informants. Time pressure and parallelism of the work prevented full use of KA results in the downstream development efforts, a source of frustration for the investigators. Some prototype problems could have been avoided had information in the dossier been more fully exploited.

4.4.3 Target Community Interests

From a stakeholder perspective, the primary role of the target community is to receive the knowledge elicited and codified by the KA effort. However, this over-simplifies both the role and the attendant demands it places on the target community:

- Being an audience for codified knowledge takes time, effort, and investment of resources. There must be sufficient incentive to the community to make this happen.
- Transfer of knowledge is more than passive "reception"; the knowledge acquired must be deployed or used effectively or the effort has not met its objectives. The real receptivity of the intended target community to the knowledge likely to be obtained must be assessed on a realistic basis.
- More often than not, the transfer community will include the sponsors and/or funders of the KA effort (since they are recipients of the results). Their stakeholder commitment, therefore, must go beyond reviewing and even utilizing results. They must buy into the value of the project up front, support it with sufficient resources over the course of the project, and, in the end, must judge that the effort expended was a good investment.

These criteria for full stakeholder commitment from the target community are by no means simple to meet. The following example from the TCIMS experience centers upon the target community, the technologists charged with building systems that operate in the focus community.

Technologists operating in the DARPA project environment experience conflicting expectations. There is a lot of pressure to construct demonstrations of varying fidelity and depth in a short time frame. The nature of the work ensures that there is an absence of consensus requirements to guide development. Further, in projects such as TCIMS, the consortium model of operation results in commercial companies each seeking viable product ideas; this dynamic can work against the competing dynamic of extensive knowledge sharing and exchange and close collaborative work. The research nature of DARPA also tends to create expectations that participants will utilize other DARPA products, that often involve unproven technology.

Motivators/Incentives

In the TCIMS environment technologists are in crucial need of usable and distilled domain knowledge. The ideal situation for the technologists would be if the final product of KA amounted to a real requirements specification. This is unlikely in practice. Other valuable input from the investigators would be tangible and innovative product ideas. This has occurred at modest levels in TCIMS due to the investigators' interactions with articulate informants capable of describing desirable future situations. Another motivational factor for technologists is that investigators may potentially relieve the technologists of the need to carry out a lot of time consuming basic research in the domain, provided information transfer from investigators to these technologists is efficient.

Barriers/Disincentives

Chief among the barriers to using the results of knowledge acquisition by the technologists is that the volume of written reports overwhelms their time and appetite for text. In TCIMS, this was a common complaint, although materials were readily available on a server. Three methods were used in TCIMS to mediate this problem: the KA was used as an "answer person" in both a review and consultation capacity; a Web-based index of the KA results and derivative products was provided; and the large body of written reports were summarized and interpreted. This was partially successful, but the fact that the KA effort was being pursued in the same time frame as when the technologist were required to build and deliver the systems prevented realization of the full benefits of these strategies.

A partial explanation of the difficulties experienced with information transfer was that many of the TCIMS KA staff were not extensively experienced in the types of formal representations which the technologists determined would be useful results. As a result, some less formal representations were constructed for their use that had the additional benefit of being more understandable by the informants. A desirable adjunct to the less formal representations would have been an automated interlingua that would facilitate the transformation of informal KA results to formal representation. (These issues are discussed more thoroughly in Section 5.0.)

Technologists work under several conflicting constraints: DARPA project goals, as well as the commercial needs of their respective organizations. KA results may impose still further, often internally contradictory constraints, or additional complexity in requirements; e.g., doctors, medics, and support personnel have varying requirements for platforms and interfaces. In the DARPA context, the KA component of the overall project may have been perceived as driven in part by its own intrinsic research objectives (as suggested in the discussion of investigator community stakeholder interests above). Some resistance on the part of technologists to receiving certain kinds of information from the KA process might therefore be predicted.

Given that some constraints will have to be suspended or at least relaxed, target community audience may be tempted to pressure investigators to resolve conflicts by imposing a normative, consolidated filter on the data. This could bias not only the knowledge transfer, but codification and the initial elicitation processes (e.g., the interview protocol). Such conflicts could surface as a stakeholder issue from the target community side (i.e., there is some data they do not necessarily *want* to know), both at the outset and at the conclusion of the project More fundamentally, it is worth remembering that technology developers may not be naturally receptive to information that implies their technology or product ideas already committed to are not likely to be utilized in the focus community. It is highly advisable to get these stakeholder interests clearly specified as early as possible in the project. Knowledge acquisition is not the same activity as focus groups for new product ideas, requirements specification and validation, or work process improvement for the focus community. Yet a KA effort may readily be characterized along these lines, and hence yield failed expectations and unused results.

This discussion has been heavily weighted to the TCIMS-relevant scenario, where the target community are technology product developers. Similar issues may arise in other cases, however.

4.4.4 Other Stakeholder Issues

We have seen that the separate stakeholder interests of focus, investigator and target communities need to be considered as part of overall KA enterprise planning. Once these interests have been identified, the relationships of specific objectives with specific interests of stakeholder groups should be documented.

It can be helpful to identify potential areas of conflict up front. Not all objectives are compatible with the same project structure. As a simple example, if the KA enterprise is viewed by technologists as a kind of pre-requirements analysis to aid in technology design, and simultaneously viewed by informants as an opportunity to air concerns about institutional practices within the organization, these competing agendas could create tensions in the overall project and at the detailed level of the individual KA session (e.g., maintaining desired focus within an interview). The choice of representations for KA workproducts is one area where such conflicts are likely to become visible quickly. To document acquired knowledge in a way immediately useful to the focus community, amenable to validation by that community, or amenable to analysis by a target community may require very different processes and representations.

If the KA enterprise goals explicitly involve organizational intervention, this should also be recognized at the outset and appropriate change management issues addressed. The KA process may bring together people who do not interact as part of routine work practice. This may create new alliances and new perspectives on organizational relations, or may stir up potential conflicts. The KA enterprise planners should consider the possibility of these dynamics up front and have strategies at hand for responding appropriately.

4.5 Selecting the Elements

Once the KA objectives for the enterprise have been established and stakeholder interests have been assessed, the key elements of the "canvas" need to be selected. Selecting the elements of KA is largely a resource optimization problem. We have a lot of ground to cover and almost always insufficient resources to do so as thoroughly as might be desired. What is the best strategy for allocating the efforts of the various investigators, informants, etc.?

For each element type, both individuals and an overall "pool" is selected as part of planning. (The "pool" of investigators might more typically be called the "KA team".) The pool concept is useful because some characteristics used as selection criteria may apply to groups rather than individuals; that is, the "pool" is more than a list of individual names. For example, perhaps only medics from a given hospital setting will be interviewed for the project. The issues constraining these choices (availability, stakeholder relations among organizations, logistics, resources etc.) will generally be at a higher level than the issues affecting choice of particular individuals to be in the pool. Thus for each of the characteristics for individual elements (described below), the pool serves as a high-level "reality check" that overall KA enterprise goals are attainable.

For each type of element, a similar kind of process can be outlined:

- Select the characteristics desired for each individual choice (e.g., all investigators should be familiar with the domain, select only informants from three key settings).
- Select the "candidate pool" that meets the overall criteria.
- Define criteria for the "pool" or set to be selected in each category. The required sampling approach will be a key element here, highly dependent on the objectives for variability data.
- Further characterize the candidate elements according to the new criteria.
- Select the elements for the pool, based on individual and pool criteria and the characterizations.
- Identify risks and mitigation strategies based on characteristics of the selected pool.

In the following sub-sections, we give guidelines for selecting and characterizing the elements of the KA effort, starting with characterizing the focus communities of practice and selecting the work settings for study, followed by the three groups that play a role in KA, the investigators, informants and artifacts, and ending with the audience and topics. Investigators and informants are chosen respectively from the investigator and focus communities of practice in the selected knowledge transfer configuration(s).

While element selections are presented in a linear fashion, the various choices are highly interdependent. In general, the best selection approach is to first make the choices most constrained by the project context, then select other elements to mitigate risks, address gaps and limitations or achieve synergies. For example, a KA enterprise could be initiated where it is predetermined that a certain group of practitioners will be the only knowledge sources directly accessible to the team. In this case, other knowledge sources (e.g., documentation, references, etc.) could be sought to create a cohesive overall set. In another case, it may be that the topics of interest are highly constrained but more strategic choices can be made with respect to the informants.

Planning for Variability

In addition to the general considerations for selecting objectives of the KA enterprise, special care needs to be taken to allow for the capture and documentation of variability. This will affect the choice of representation of knowledge in the project. It will also have an impact on planning the various "threads" so that varying information will be available in a single setting, to allow for modeling of the variability.

Potential barriers to capturing adequate models of variability should be noted as risks in the KA plan and strategies put in place to address these barriers. These obstacles might include the following: pressure on the part of practitioners to project a "normative" view of their work practice; difficulty on the part of investigators in representing the variability with adequate traceability back to distinct settings, informants, etc.; and unwillingness of technologists to work with data on variability that does not fit their model of what can easily be accommodated in implementations.

Eliciting information for systematic documentation of variability may require specific techniques within the KA repertoire. These could include techniques for extracting variability data from workproducts such as interview reports or videos, where there was a different primary purpose of the data-gathering task. Viewing a number of artifacts or KA workproducts in sequence, where there is a clear basis for comparison, is a quintessential technique for eliciting variability information. The clearest way to articulate variability is either side-by-side or successive viewing of multiple exemplars exhibiting common and variant features. This allows human pattern-matching cognitive skills to be applied. The knowledge acquisition plan might need to allow for collecting data from multiple settings, elicitation from multiple informants simultaneously, or analysis of similar or analogous artifacts "side by side."

> Example. Suppose an investigator reads an interview report with a medical rescue helicopter pilot, then reads a second report with another pilot from a different facility. Provided that the purposes of the two interviews were similar and the performance of the data gathering session similar in nature (ideally, with similar questions asked) the investigator could begin to note variances and discrepancies in terminology, procedures described, etc. Here, it is precisely the placing of the separate workproducts in a common interpretive context that provides the basis for eliciting observations of common and variant aspects of the data.

4.5.1 Characterizing Communities of Practice

The communities of practice for the KA enterprise were essentially selected in establishing the overall objectives, and stakeholder issues for these various communities have been considered in some detail (as discussed in Section 4.4). As a precursor to selecting specific settings and informants for study, it is helpful to consider overall characteristics of the focus communities involved. For each of the following characteristics, we provide some hints about the impact it can have on the knowledge acquisition process. The discussion uses the term "field" as a rough synonym for community of practice, because typically the professional aspects are established for a much broader level than the individual community to be studied. For example, while surgeons at a particular hospital may have some "local culture" there are important factors to be considered about the field of surgery as a whole.

Nature of expertise in the field

Is the field process-intensive or performer-intensive? In very mature fields, accepted practice is so formalized that it has been codified into a rather rigid process. A prime example of such a field is civil engineering, with a base of knowledge of practice several thousand years old, and where the cost of mistakes is such that process formalization exists in part to reduce risk. Additionally, in civil engineering, the behavior of the materials and forces that affect structures is relatively well understood, both due to hard-won experience and due to the development of analytical models that have some predictive power. This is not to say that performance in such fields does not have a subjective aspect. In civil engineering, the subjectivity is the artistic license that the engineer exercises in making an aesthetically pleasing design. The maturity of the field makes it possible to have a clear distinction between what is subjective and what is not. For the objective part, performance can be judged by how well the process is followed; legal liability may even be associated with failure to follow the process. Such fields are characterized as *process-intensive.*

Medicine, in contrast to civil engineering, could be characterized as *performer-intensive*. Although the practice of medicine has a similar long history, the repeatability of results is drastically less than in civil engineering, in no small part due to the complexity of human physiology, the variability of effects of treatment from individual to individual, and the changing cultural norms and accepted treatment practices. As a result, there is an accepted body of practice that leaves much discretion to the practitioner working within the guidelines of that accepted practice. In such a field, it is difficult to determine when a particular performance should be judged on objective grounds, and when it is part of the subjective skill of the practitioner. In medicine, the subjectivity lies in the professional discretion of the diagnostician.

Communities in process-intensive fields are more likely to be interested in an accepted, consensus view, while in performer-intensive fields, careful study of variability is likely to be useful. KA in process-intensive fields is more likely to encounter resistance based on fears on the part of the practitioners that their positions might be jeopardized, if the KA project is successful.

Nature of training in the field

The general approach to training in a given domain has a direct impact on the approaches one takes to deriving information from the informant pool. Disciplines in which analysis and process are highly prized usually offer informants who are trained to reflect upon their practice in detail and can offer well-structured accounts of their work. By contrast, trauma medics are trained to carry out procedures in response to recognized situations. The time critical nature of the trauma domain makes it very important for the practitioner to assess a situation, match the pattern of the situation to the procedures needed, and act with dispatch, else the well-being of the casualty or injury may be irrevocably compromised.

A similar distinction is reported by Brigitte Jordan as a result of fieldwork on the effectiveness of training programs for traditional birth attendants in the Yucatan [18] funded by industrialized nations. Her research suggests that the low effectiveness of much of this training results from misapplication of didactic modes of teaching in cultural situations where learning by an apprenticeship mode is more appropriate and culturally familiar. A training style based on Western norms of professional medical practice typically makes heavy use of sequential scenarios that move through the various steps involved in a medical procedure such as assisting at a birth.

In contrast, an apprenticeship-based model might categorize tasks according to criteria such as the degree of skill, personal authority and experience required to carry out the tasks. Such a model would thus approach an overall scene description in concentric circles of peripheral support tasks (suitable for performance by a less experienced apprentice) and core tasks to be performed by the expert. Jordan was initially surprised to find the traditional birth attendants reluctant to share stories, which correspond to what is familiar in Western medicine as "case studies", a standard way for physicians to communicate both among themselves and with other communities. Hence even the ability to elicit knowledge from an informant via a sequential, hypothetical scenario already presumes a great deal of cultural context that might not be applicable in other domains or other cultural settings.

Stable versus fluid domain knowledge

The state of knowledge in a given field has a major impact upon the approach to utilizing informants or artifacts in the field. A domain might contain considerable amounts of stable knowledge. For example, in the civil engineering domain, there is a substantial body of settled knowledge relating to aggregate experience about constructing common structures within bounds of size and characteristics of local geology, given the parameters of the building materials. If some part of a domain is undergoing rapid change due to immaturity or high-impact advances in technology, the "half-life" of knowledge for new practitioners who deal with that part of the domain is shorter. In civil engineering, innovations in composite materials make the half-life of any knowledge based on the use of known traditional materials very short. Medical procedures in trauma care and emerging fields such as personal communications systems also have this fluid quality.

The overall objectives for the KA effort need to be appraised for realism relative to the rapidity of change in the field of interest. Domains that incorporate dynamic knowledge will require some more stringent techniques for cross-validation. Strategies used in the TCIMS project included combining interviewing with observation or cross-checking data from experts with knowledge of varying degrees of outdatedness.

Stage in business lifecycle

Businesses often go through a lifecycle that includes, roughly speaking, a start-up phase where the knowledge is volatile and innovative, a maturity phase where the knowledge is stable and well-accepted, and a decline phase, where the knowledge becomes obsolete and irrelevant. There is a trade-off of the effectiveness of KA against the need to do it in each of these phases.

During the innovation stage, the need for a well-organized view of the knowledge in the field is high, since it allows technology planners to see the directions that the field can take. Unfortunately, it is also during this time that the value of the KA results are short-lived, and the time taken to complete a systematic KA process might outlive the usefulness of the knowledge.

At the other extreme, when knowledge is stable in a mature business, the need for organizing the knowledge is less, since the field has managed to do some of this organization in its own practice. The effectiveness of KA is, however, high; since the knowledge is well-accepted and stable, it is

possible to get a solid, coherent corpus of information about the work practice. During decline, the need for knowledge acquisition might rise again, as the number of working practitioners declines, and the knowledge is in danger of dying out.

A setting for knowledge acquisition should ideally be chosen from a business whose position in its lifecycle balances this trade-off, so that the KA is effective enough to provide value, and the need is high enough to appreciate this value.

Degree of professional stratification within community

Unless the focus community is rather uniform, such as professors of philosophy or long-haul truck drivers, some degree of professional stratification is likely to be observed within a focus community. The domain of medical practice is highly stratified, both in training and legal certification of the practitioners. Physicians have traditionally considered themselves to be the central focus of the medical field (although this is certainly affected by the emergence of HMOs.) At the same time the experience in TCIMS is that while physicians are usually the prime source of information on treatment and diagnosis, they have less detailed knowledge of the enterprise of medicine than other practitioners. A compelling example is that nurses are more knowledgeable about the actual workings of hospitals than physicians because of the division of labor that exists between nurses and physicians in most hospitals.

This state is not unique to medicine. It is common knowledge that the secretarial staff in most organizations has more immediate knowledge about the real workings of the businesses in which they function than the acknowledged decision makers. As a result it is suggested that the focus community be covered as completely as possible by the KA plan, including subgroups that might be considered as "non-professionals" to ensure that important features of the domain are not obscured or misrepresented by the KA products.

4.5.2 Selecting and Characterizing Settings

One of the most important decisions to be made in planning the KA enterprise is to determine the settings that are to be studied. Settings are specific work environments where the knowledge of a community of practice is created and/or used. The choice of settings will affect the informants selected, the types of knowledge that can readily be obtained, the type of elicitation session (i.e., interview, observation, instrumented data logging) most relevant and feasible, and other factors. Basic considerations for selecting settings include the following:

- For knowledge acquisition in technology-intensive environments, a variety of potential settings should be considered (end-user operational setting, application development setting, maintainers' setting, etc.). Linkages between settings should be considered (for example, the developers' and users' settings for a particular system).
- The overall variability objectives for the KA effort will affect how many and which settings are selected. Which aspects need to be documented? Informants in different settings will often use a different terminology, which might overlap with other terminology in misleading ways.
- Some focus community settings might also be target community settings, in which case the informants selected may be part of the eventual audience for the knowledge gathered. While this overlap is generally beneficial, it may be advisable to arrange that some of the intended audience will have had not prior contact with the KA team, to make sure that knowledge has been codified in an effectively transferable way.

Settings chosen for study may involve the overlap of several distinct communities of practice. Not all of these will necessarily be within the scope of the project or singled out as a source for informants. Working from the settings selected, and the communities of practice identified in the enterprise objectives, it is possible to identify "pools" of potential informants with relevant experience or roles in a selected setting and membership in a relevant community of practice.

> Example. Surgeons, medics, nurses and paramedics all draw on similar training for their terminology; therefore, as far as terminology is concerned, distinguishing between these three groups might not be considered relevant for a KA effort. However, among the on-site trauma personnel, there are also medical technicians, transport personnel (ambulance drivers, helicopter pilots) and other support personnel (local fire fighters, policemen). If any of these will be involved in the knowledge acquisition effort, then it is appropriate to identify them as separate communities of practice, based on differences in terminology. Among the medical personnel, the difference in professional status and responsibilities indicates that nurses and surgeons might be considered as separate communities, if the goals of the effort include information that is embedded in this distinction.

4.5.3 Selecting and Characterizing Investigators

Depending on the structure of the project, there might be a possibility for selecting the pool of specific investigators from the investigator communities designated in the knowledge transfer configuration(s) for the KA enterprise. Even in situations where choice of investigators is fairly constrained, it is still worthwhile to characterize the investigators from the standpoint of both specific knowledge and capabilities. This will facilitate selection (where this is possible), teaming and partnering choices, decisions about which investigators will perform which sessions, identification of potential risks and, for larger-scale projects or more permanent teams, development of a training or capability development plan.

The following questions can be used to characterize the relevant knowledge of the investigators:

- *What is the investigator's level of familiarity with topics of anticipated relevance to the project focus?*

  The knowledge (or lack of knowledge) an investigator brings to a session is a two-edged sword An investigator already familiar with the field of study, or with a closely allied, has a clear advantage in some situations. The investigator's knowledge of a domain can enable them to ask good probing questions, to recognize accurate data and be cautious of questionable data, and can establish credibility with informants. This knowledge can be particularly crucial for artifact analysis, since it can be quite daunting to try to understand an advanced document without familiarity with the basic terminology of the field. Certain artifacts might be incomprehensible to an investigator without appropriate background.

  This can also be a relevant factor in interviewing experts with very high demands on their time. Such informants may be offended to be interviewed by someone with insufficient background. If they need to spend time explaining fundamentals of the field, or material available from other sources, they may justifiably feel that their time is not being well spent and lose confidence in the KA process.

  Yet investigators' knowledge of the domain can also be a barrier in certain circumstances. It can tempt investigators to "be their own informants" and impose their own domain knowledge on their informants, or otherwise bias the write-up of the acquisition session. Without formal validation of resulting KA workproduct(s) by the original informants possibilities for such confusion become stronger.

Even for investigators scrupulous about keeping their own views separate and partitioned from the data derived from their informants, the influence of previously held knowledge may affect what they do and do not ask and the terminology they use in writing up the session. Over-familiarity with a field can blind the investigator to unarticulated knowledge embedded in the community of practice. Given that eliciting such embedded knowledge is often the most elusive goal of a KA project, the KA effort should also, by design, include some investigators naive with respect to the focus community. Such investigators should be well-trained in KA techniques and principles, however, so that they can take advantage of their naivete; otherwise, they are simply under-informed investigators.

- *From which community was the investigator drawn?*

  This question was raised earlier, in Section 4.4.2, in discussing investigator community interests. Here the question is relevant to assess what knowledge base the investigator will have through prior experience. It is largely relevant to the previous question; in addition, it can be advantageous for the investigator to understand the organizational context of the setting where KA is being performed (i.e., the stakeholder picture).

- *Does the investigator play a "bridger" role across more than one community? Does the investigator have any previous experience in playing a "bridger" role between communities?*

  Since investigators often become ambassadors of sorts between focus and target communities, they should be able to understand motivations from both points of view, and speak a language comprehensible to both communities as well. This ability also relies on the personal interaction skill of being able to "switch hats", and fit in with multiple communities.

- *What is the investigator's level of familiarity with various knowledge representations? What are the investigator's skills in artifact analysis?*

  In a particular KA project, a large number of different representations might be used for the KA workproducts. Some of these representations will be intended for feedback to the informants, while others might be of use to system developers, and others might be for internal record keeping by the KA investigator team itself. Some of these representations can be difficult to use, and fluency with them is a skill that could discriminate one investigator from another. Some sample representations, and guidelines for their use, are given in Section 5.0.

- *What are the investigator's preferred modes of communication?*

  (e.g., extended telecons, face-to-face interviews, one-on-one vs. group meetings, email, written correspondence)

This is a "starter set" of questions for characterizing investigators for the purpose of selection and thread planning. Investigators are selected and characterized based on their sets of skills and degree of maturity in applying those skills; this information will be useful in planning the investigator's thread throughout the knowledge acquisition project.

A more detailed set of questions can be applied in assessing general skills and capabilities for KA. These can form the basis of an ongoing assessment and capability development plan for the investigator team, discussed in more detail in Section 4.10.

4.5.4 Selecting and Characterizing Informants

Informants will be members of the focus communities in the selected knowledge transfer configurations. The Canvas framework encourages consideration of informants from a number of settings, and at varying levels of training or skill. There are many factors to take into account in choosing a pool of informants from the focus community. These factors include characteristics of the overall focus community, particular sub-communities, groups and settings, and characteristics of individuals.

The selection process for informants is involved enough that it is advisable to get help from knowledgeable people in the community of practice at the outset (essentially, informants about who to select as informants). In TCIMS, a body of experts in the medical domain were heavily relied upon for guidance and selection of high-value informants.

The following paragraphs outline some broad factors to be considered.

Breadth of experience

An informant who has played several roles in the work setting, and has seen several changes in the work setting usually has a unique insight into how the setting works. This is not the same as depth of knowledge; even a casual user of many different processing systems has quite a sophisticated idea of what they can and cannot do, and can be a great source of comparative data, even if the user has never learned enough about any of the systems to customize them.

Obsolescence of domain knowledge

While it may seem obvious that one should avoid selection of informants whose lack of current involvement makes their input suspect, this is not altogether adequate as a criterion. For example, there may be a trade-off in the availability of informants with the degree to which their knowledge is up-to-date. Also, depending on the stability of the knowledge in the field, being up-to-date might not always be the topmost concern.

Some care should be taken in complex domains to allow informants to validate their own input. For example, the intense field experience of knowledgeable combat medics in the U.S. military now dates from the Vietnam war era. In this case it is necessary that medics be given the opportunity to provide their recollections of that experience, review draft session reports and note where current practice diverges from that reported. Further, the aggregate session reports may be examined by the aforementioned body of experts in the domain for a more global validation. In this case, not only can experts with less current knowledge do some of their own validation, but new knowledge may be created by making the difference between older and more current practice more dramatically visible.

An extreme example along these lines is the experience of former air traffic controllers, brought back after lifting of the long-standing ban on rehiring, described in [13]. In this case, the returning controllers were aghast at the pressure and chaos of current working conditions compared to the conditions that, more than a decade earlier, were considered severe enough to motivate a general strike in which safety issues were a major area of concern. Although these more experienced controllers probably have a wealth of experience that could be tapped, by and large they cannot adapt to new working conditions. This is a case where knowledge acquisition techniques could be used to derive the particular historical perspective of the veteran controllers.

Ability to articulate knowledge of actual practice

Some of the most valuable informants in a domain are those able to analytically describe *actual* practice (as opposed to official or stereotyped reports of practice). In some cases this is difficult to achieve due to lack of ability to articulate on the part of informants. In other cases it may reflect the nature of the performance patterns of the informant community. In the case of Trauma medics, it was discovered that there was a lot of value in performing field observation of medics in action to cross-check and elaborate on the basic interviews. Because of the split-second pace of decision making in the field, it is difficult for trauma medics to fully describe what they do, as much of their behavior is learned reaction, and difficult to describe outside the practice setting and to those unfamiliar with the domain. TCIMS also derived benefit from access to practitioners who were qualified instructors.

Reflectiveness /Introspectiveness

The ability to reflect upon one's expertise and analytically report the results of that reflection is affected both by the nature of the domain, the personality of the informant, and the structure of the training in the field. The experience in TCIMS indicated that trauma medics are not ordinarily introspective about their practice unless they also happened to be instructors. The hypothesis offered is that this is largely a side effect of the time-critical nature of their work as well as the training to react quickly to situations.

Other traits of "attitude" that can make an informant a good participant, particularly in a collaborative approach to knowledge acquisition, include the following:

- Ability to admit to lack of knowledge: it is helpful to have informants who can be forthright about when the session has moved to a boundary of their own expertise that significantly changes the accuracy or quality of their data. Posturing (pretending to know more than they do) or, conversely, reacting over-apologetically to being asked questions outside their area of confidence, are both problematic reactions.

- Tolerance for investigator's level of knowledge: Although a good plan should ensure reasonably well-prepared investigators, there is still a lot of patience required in stating the obvious for long periods of time! Once again, there can be a variety of reactions that are not particularly useful to the outcome, including: getting insulted; getting patronizing and turning the session into a tutorial (mistaking the learning of the investigator as the goal).

4.5.5 Selecting and Characterizing Artifacts

Some issues in selecting an artifact for study parallel those involved in characterizing informants; e.g., issues of currency of knowledge (though an artifact might not be as insulted to be considered out-of-date as a human informant). Since artifacts are workproducts in their own work setting, they share some features in common with the workproducts of the KA setting itself.

Certain factors such as the accuracy and quality, and currency of materials will be of obvious importance in selecting and characterizing knowledge sources. Here we outline some additional characteristics of artifacts particularly relevant to the selection process.

Accessibility

For humans, accessibility usually has to do with the demands on the human's time. In the case of artifacts, accessibility is often a matter of politics. Some artifacts might be protected by commercial agreements (such as consortium confidentiality), while others might be classified by a

national government (often the case in military projects). Some artifacts will cost money (e.g., the executable code for a commercial system), which might be beyond the project's budget. Some artifacts (e.g., patients' medical records) will also raise privacy issues.

In addition to these questions of access in terms of data rights, there are a more pragmatic set of accessibility issues, such as availability of data in on-line form; and whether the format of the documents will facilitate the selective, filtered examination that is usually required in KA.

Intended audience

As is the case with knowledge acquisition workproducts, artifacts typically have an intended audience, which affects the representation used, the terminology, and even the accuracy of the statements. For example, a number of commercial systems have "beginner's guides" that do not reflect all the complexity inherent in the system. Some beginner's guides might even present factually incorrect simplifications for the purpose of protecting a naive audience from overwhelming complexity. Evaluating the context in which an artifact was intended will usually require some *context recovery* (i.e., identifying and making explicit the assumptions embedded within the software artifacts that come from the cultural context in which they are used.)

Interpretativeness

Artifacts can display varying degrees of interpretiveness or reflectiveness in their content. In a system engineering context., for example, some artifacts relate directly to a single system, and do not attempt any degree of generality about what they do as a member of a class of capabilities. Examples of such artifacts are user's guides to software packages, or executable code itself. Other artifacts explicitly try to place the system they describe into a larger context, either by comparing it with other similar systems or describing it in more general terms. Examples of this sort of artifact are survey articles written by focus practitioners and translation packages that adapt the interface of one system to the standards of another.

It is not necessary for an artifact to refer to several systems to be interpretive. Rationale documents may refer to a single system yet attempt to document why the system was built the way it was. Such documents typically relate a system to some broader design concepts, and may predict how the design will behave when it is adapted to new uses.

Depending on the degree of interpretiveness of the artifact, it will be important to understand the context in which it was created, the range of experience of the author, and the intended audience. This is because the raw data in such an artifact may be closer to the eventual representation desired for the target audience; yet the data may represent the un-validated and biased view of a single informant (and, of course, stakeholder) from the focus community. It can actually be more difficult to make appropriate use of such materials, therefore, without a good deal of supporting context.

4.5.6 Characterizing Audience

The audience for the knowledge acquired is the target communities in the selected knowledge transfer configuration. The audience for knowledge acquisition is usually chosen when the KA enterprise is initiated. Even if the audience has been chosen, it is still necessary to characterize the audience to esure that the knowledge gained matches audience needs. Audience characterization is also used to ensure that representations used to transfer knowledge to the audience are suitable for the audience.

Questions for direct inquiry include the following:

- What do you know about the topics of interest to this project?
- How sure are you of your knowledge?
- What do you need to know?
- In what form do you need the information?
- (Key question) How will you use the knowledge once you have it?

Other questions for assessing the audience include the following:

- How open are they to new learning?
- How busy are they? What will they have time to absorb in order to make practical use of the knowledge acquired on the project?
- What is their familiarity with various representations? This question must also be addressed for those members of the focus community that will be validating the data acquired.

4.5.7 Selecting and Characterizing Topics

Topics should be chosen for each phase of knowledge acquisition and for each thread and session within a phase. If the investigators do not have knowledge of the KA domain up front, it will not be possible to determine topics before knowledge acquisition begins. If investigators do have knowledge of the KA domain up front and choose topics for investigation, it is possible that investigators may be biased based on their current knowledge. See Section 3.2.3 for more discussion on investigator bias.

Topics will evolve as more knowledge is gained of the domain. These topics should be included as part of the models in the domain dossier. The topics will likely fall into a hierarchical relationships, which is often modeled using taxonomic modeling, as discussed in Section 5.6.4. Topic selection should be a continuous process during knowledge acquisition as more information is gained and modeled.

4.6 Selecting Representations

One of the most important planning steps involves selecting the set of representations that will be used to codify the knowledge elicited in various sessions.

The idea that representations need to be selected by intention may surprise people who are used to working from the standpoint of a particular software development methodology. Most methods include as a central part of their structure a preferred set of notations or representations to be used. Often these are correlated with specific steps in a development life cycle. Adopters of a method generally work from the premise that this set of representations is the best choice for those types of problems to which the method is targeted. Such sets of representations also facilitate clear semantics and relationships among representations, exchange of data across projects in standard formats, and standardization via tool support.

Of course, developers know the difficulties in applying such representations in practice. The technology user community may be unfamiliar with the representations; this raises problems when it

is desired that some workproducts be collaboratively created or validated by users. Even developer groups will have varying degrees of difficulty in adopting (and may even resist adopting) particular representations, or will apply and interpret them incorrectly.

From the Canvas standpoint, part of the difficulty in such situations is that method developers and adopters may be unaware of the extent to which their preferred notations represent interventions in the various communities involved. Although this does not directly challenge the value of methods it does suggest that knowledge acquisition planning must offer a systematic way of considering the impact of these interventions. In Canvas, the notion is that competent knowledge acquisition requires the availability of a ***repertoire*** of representations. The repertoire concept implies that no single technique is appropriate for all situations. Instead, a selection process is required to suit a given technique to a given situation. Each representation facilitates certain kinds of knowledge codification and discourages others; each has inherent biases.

In Section 5.0, we explain the constituent elements of representations, catalogue a number of different representations and describe some of the uses for which each is particularly suited and ways in which they can be combined. Here, we describe the specifics of how the notion of a repertoire of representations is utilized in the planning process.

Within any community of practice, certain representations will be preferred ways of capturing and exchanging knowledge. The distinct set of representations used by that community can be considered as the repertoire for that community. To the extent that these form part of the ordinary work flow for the community, there may be no necessity for members to be overly conscious of the biases of their representations; they can remain implicit to some degree. (In these terms, a software development methodology, as discussed above, provides a systematized repertoire of representations.)

However, since a primary goal of KA (in the Canvas context) is to effect transfer of knowledge *across* communities of practice, KA planners must take a more informed approach to selecting representations. Specifically, for each stakeholder community (as determined in Section 4.5.2) it is important to know which types of representations are part of the repertoire for that community. This will provide a basis for deciding which notations are appropriate for capturing knowledge elicited by investigators and passing that knowledge on to the audience.

Different representations will be best suited to each community of practice. For example, certain representations that are familiar to computer scientists (e.g., flow charts) might be unfamiliar to medical personnel. Other representations might well be familiar (e.g., hierarchical tree diagrams), but have a different meaning. The representations in use by each specific community will have different importance depending on the KA role of that community:

- *Representations of the informant community* will affect the kinds of artifacts that may be available for study by investigators. (This may require investigators familiar with these representations or training to create familiarization.) The informant community repertoire will also guide the selection of representations to use in collaborative modeling sessions, where it is desired that informants participate directly in creating KA workproducts. Representations used for validation sessions will also be determined, and the potential direct value to the focus community of Representations created for other audiences can also be determined.

- *Representations of the investigator community* (if this is a distinct community) will help determine how well this community can play the "bridging" role of facilitating transfer from focus to target community. In addition, awareness of the Representation repertoire of the investigators (even on an individual basis) is an important element in identifying potential sources of bias.

- *Representations of the target community* will, of course, determine the best ultimate form of codification to meet the overall objectives of the KA enterprise.

In selecting from available representations for various uses during knowledge acquisition, consider the following:

- Ability of the investigator to facilitate knowledge acquisition sessions that use the representation.
- Ability of the audience to understand the representation
- Appropriateness of the representation for the goals of the KA enterprise
- The amount of effort it will take to translate from one representation to another if multiple representations are used. Are any tools available to check the consistency of information captured in multiple representations?

A catalogue of representations to be used in the project, and the audiences for which they are intended, will make planning sessions much easier. In TCIMS, information of this sort was encoded in a template for a session report, with boxes to be checked for each representation used.

Consider that there may be more issues than mere representational competence in matching representations to various communities. There may be stakeholder issues involved in selecting various representations. A group might reject a representation even if it is simpler than that to which they are accustomed (or perhaps *because* it is simpler). representations may be associated with other communities, and thus representation selection may reveal other stakeholder issues. representations have strategic as well as technical import that must be considered.

4.7 Initializing Dossier Infrastructure

All of the information determined at the enterprise level can be used as an index for the dossier of the workproducts to be produced by the KA project. Section 6.0 gives a detailed description of how one can build a dossier, and starter sets based on the enterprise planning components for building an index. All material (or references to the material) are kept in the dossier, including:

- Materials read or developed in preparing for a KA session;
- Session write-ups; and
- Representations of knowledge acquired in the chosen representations.

Materials in the dossier can be indexed by a variety of characteristics, including:

- Topic
- Source
- Audience
- Representation used

There is always some informal KA that has taken place prior to doing the formal planning process. The informal data gathered so far should be used to "seed" the dossier once the structure is set up. This serves to validate that the dossier structure is adequate, and gets the data accessible to the entire KA team.

4.8 Planning a Thread

The notion of threads was introduced in Section 3.2.3. In that section, we looked at the interaction of each type of thread with a particular session. In planning a thread as a whole, we must view the *sequence of sessions* that a given person (investigator, informant) participates in over the lifetime of the KA effort, or that affect the incremental interpretation of a given artifact or workproduct. In this section, we will discuss issues to consider in planning the various threads of the Canvas framework.

Thread planning is the most challenging planning task in the Canvas approach, because it bears least resemblance to conventional project planning, and because it is where the need for adaptive and dynamic re-planning of the KA enterprise comes to the fore. The term "planning" might imply an up-front, detailed plan that outlines how the entire project will unfold; in fact, the notion of threads helps make clear why any KA plan can provide only a starting structure.

After each session, new knowledge sources may be revealed that need to be considered in future thread planning. Sessions may yield more, or less, knowledge than anticipated; certain topics may be revealed as dead ends, while others, unknown at the start of the effort, move into the foreground. Just as a loom can be set after each woof strand to form a pattern, the KA enterprise is set after each session, to take into account the effect it had on all threads. Thread planning, therefore, involves determining which aspects of an element are to be tracked, so that further planning can be performed based on the results of sessions. Automated support would make it much easier to envision the manifold cross-connections between elements (investigators, solo and in partnership, informants, artifacts, etc.) and sessions over the lifetime of the KA enterprise.

As the project progresses, new combinations of expertise brought together by the KA effort may spark surprising and serendipitous discoveries. Individuals may reveal unanticipated strengths or stakeholder issues (e.g., an informant turns out to have a "pet theory" about his domain that has never been published, but when documented in the KA context, wins buy-in from other informants). Such events cannot be planned for, but can be anticipated by carefully tracking the assumptions guiding each thread's development.

Thread planning includes a number of factors including the overall development life cycle for the participant (or other element of the KA process), the sequencing of different kinds of sessions, and the timing of events along the thread. Here we provide a few suggestions of issues to consider in thread planning for the major elements, investigators, informants and artifacts. The overall life cycle of a thread for a human agent generally starts with orientation and familiarization sessions, then traces a path through various topics, representations, types of elicitation sessions.

Investigator Thread Planning

At the broadest level, the investigator thread may include learning goals as well as goals that directly and immediately contribute to the project. This may include rotating investigators to cover multiple topics or settings if breadth is desired, or keeping investigators focused on particular areas if it is important to have people with solid context established for subsequent sessions. These are tradeoffs that affect planning at the level of the investigator team as a whole, not just an individual investigator.

Scheduling and Rhythm of Sessions

One of the worst problems identified in our case studies is the simple problem of adequate preparation, debrief and documentation time. It seems that the less technical nature of KA activity dooms it, when done in a technology-oriented environment, to continual under-estimation of the

effort required to codify the knowledge obtained. Expending the resources to do knowledge elicitation, and not to allow ample time for codification, is worse than wasteful. Informants' time is a finite and scarce resource; the interview that does not get adequately documented literally may not happen again.

The problem of time affected when preparatory materials were received prior to sessions, the amount of time budgeted to review those materials, write-up time after the session. With respect to elapsed time between sessions for the investigator, our limited pilot experience suggests that there is a trade-off to consider. Given adequate time to document the previous session, there appears to be an advantage in a certain amount of "slack" in the schedule (up to about one week) during which time fresh insights can emerge from reflection on the session. Delays much beyond this, however, affect continuity of context and reduce effectiveness.

Depth vs. breadth of domain coverage

In determining how to allocate investigators to topic areas, one important trade-off involves focusing an investigator on a particular topic area, developing his/her knowledge over a series of related sessions, or rotating them to cover different areas. At a different level of planning, this would also apply to an investigator's assignment to different work settings, informant groups, etc. This latter trade-off is closely connected with the variability objectives for the project, since comparative insights are fostered by exposure to multiple settings and analysis of their common and variant features.

> Example. On a domain engineering project which involved the study of several legacy systems, one investigator has the opportunity to reverse engineer architecture diagrams for several different systems in succession. As a result of seeing the different design approaches in succession he is able to identify some common features of the systems that might not have drawn attention in studying a single system.
>
> In the TCIMS project, similarities and differences between trauma care in a military field setting, urban emergency room and rural settings

Bias management

Bias is perhaps the most important aspect of a participant in the KA process to be managed. The most important biases are those of the investigator, which will cloud the information that he acquires, either from artifacts or through interviews with informants.

Bias can be controlled by taking into account what information a given KA participant has received at a given point in the participant's thread through the KA enterprise.

> Example. In the TCIMS project, videotapes were made of the interview sessions. After the session, the information was written up in a report, which would go into the dossier. The reports could well be written by people who were not present at the interview itself, by viewing the video.
>
> The question of bias arises when we decide whether the person who was present at the session should brief the other team members before they watch the film. On the one hand, the briefing could clarify issues that are not on the tape, such as whether the informant was interrupted from another task when the investigator arrived, or whether he was waiting, apparently idle. On the other hand, the briefing could bias the viewer towards certain interpretations of the events, causing him to miss others. A similar problem arises when deciding whether the film should be watched in a group or separately.

This trade-off can be generalized to any KA workproduct setting; in principle, the investigators can examine workproducts in parallel or in sequence, and can consult one another's commentaries or not. Consulting earlier interpretations can save time and allow for deeper study of detail, while parallel uniformed viewing results in a broader range of interpretations.

Investigator as "surrogate informant"

It is often the case that the investigator comes to know so much about the focus community, that members of the target community will bring questions to him, rather than go to members of the focus community. The KA plan can include intentional development of one or more investigators to play this role; in this case, the investigator should either begin with a familiarity with the focus community, or should be involved in KA sessions in which members of the focus community are treated as knowledge sources. This strategy has the advantage that the answer person has access to knowledge that was never codified. It has the disadvantage that it discourages comprehensive codification of the knowledge, thereby weakening the value of the dossier and risking that the KA effort is converted into the personal learning of the investigator as "surrogate" informant.

Designing training into the process

Investigators begin relatively naive and become more and more informed about the focus community as the project moves on. This can be used to the advantage of the project, as a means of bringing in new investigators after the project has begun. New investigators can be used for sessions in which it is deemed advantageous to have an investigator who is unfamiliar with the domain, while the more experienced investigators are used for those sessions where familiarity is called for. partnering veteran with novice investigators is a classic way to obtain the benefits of both viewpoints in a session while also helping to orient the new investigator in an "apprentice" role.

Resource management

As the threads interact in a large project, there may be problems of resource management. Some investigators will have developed the background to examine certain artifacts or understand certain informants, while others will not. The management of a thread will have to take these resource allocation problems into account, and perhaps develop a thread in a particular way so as to minimize bottlenecks later on in the process.

Informant Thread Planning

Planning an informant's thread also involves managing several different learning life cycles. For most projects and most informants, there may be only one or two sessions involved in KA. However, other projects, particularly those involving in-depth knowledge capture, might require a number of sessions over time. These must be managed carefully from the standpoint of utilizing the informant's time effectively and respectfully. The more involvement with the KA effort a single informant will have, the more an additional learning cycle must be considered: that is, the informant's gradual familiarization with the goals and methods of the KA team. The following paragraphs highlight some of the major issues to be addressed:

Repetition (asking informant same question twice)

Investigators can make nuisances of themselves if they continually ask the same informant for the same information, or ask for inappropriate information. The means of managing this aspect of the informant thread is to keep careful track of what information has been gathered from which informant, and to consult this information in preparing for a session with that informant. Repetition

management may conflict with bias management, since it might be deemed appropriate for the investigator to remain ignorant of some result of the previous interview, for purposes of controlling bias.

Interview consolidation

Closely related to the management of repetition is the management of the overall interview process. If several investigators are interested in related topics, all of which are available from a single informant, it is in the interests of the project to consolidate these interviews. This requires a similar infrastructure as management of repetition, only for the information requirements of the investigators as well as the results.

Progression in representations

In scheduling multiple sessions with a given informant, one design goal is to let each session build naturally from the insights of previous sessions to elicit more detailed knowledge, closer to the form desired for the target audience. One session might involve a descriptive walk-through of a practitioner's daily workflow; from this session an investigator might identify several areas to pursue in more detail. This progression could be a combination of gradual elaboration of the topic data itself, as well as the informant's growing level of confidence in working with more formal representations.

Descriptive to innovative elicitation

The ODM domain modeling process model involves a gradual shift from purely descriptive modeling, to comparative, to evaluative, and finally to innovative exploration of new possibilities. This progression can be a useful framework to keep in mind when planning an informant's thread. The eventual goal is to develop the relationship with the informant towards increasingly collaborative knowledge elicitation and, ideally, knowledge creation.

> Example. Using KA in a requirements gathering context, this thread planning approach might involve beginning with description of current work practices and legacy system interactions, then involve informants in some group sessions where they encounter people performing similar tasks in different work settings. This creates opportunities for comparative information gathering and capture of underlying rationale; i.e., "why *do* we do it this way?" This creates the dual advantages of grounding the process in rich descriptive detail but thinking in terms of alternatives and new possibilities. A final, innovative stage might involve collaborative envisioning of possible new system capabilities.

Fostering informant reflectiveness

As informants move through the KA process, they will be encouraged to think about their work practice in new ways, and will be exposed to possibilities afforded by a knowledge acquisition effort for codifying, analyzing, summarizing and comparing information. Informants so inclined may become more reflective as this process continues, thereby changing their categorization according to Section 4.5.4.

Planning Other Threads

We have highlighted issues for two of the most important threads in the Canvas framework. Similar issues need to be considered for the other threads as well; these are not discussed to the same degree of detail here. Issues to consider include the following:

- *Artifact Threads:* How many people should review this document? In what order should they review it? What filter should each apply on the information to glean from the artifact? Should they view each other's comments and annotations? The annotations and comments are an indirect form of communication among investigators.

- *Topic Threads:* In what order should focus be "rolled out" on different topics or domains? Do certain topics serve as an introductory foundation for others? Can some be addressed in parallel, by different investigators? In progress on a given topic or domain, should artifacts be studied, followed by interviews?

  At a broad level, this level of thread planning may coincide with the phase planning discussed in Section 4.3.4. For example, on TCIMS, military, civilian urban and civilian rural domains respectively were investigated by the team.

- *Setting Threads:* What will be the sequence of KA sessions occurring as events within a given setting? How much observation will be done? Will the KA activities disrupt the work setting activities or change the dynamics enough to effect the data?

The purpose in this section has been primarily to show how the Canvas framework structures the planning decisions that must be made, so that the potential impact of each decision on the many interdependent elements can be assessed. We have seen that KA planning involves decisions at the level of the KA enterprise as a whole, at the phase level, and in the planning of the threads for each key element selected. Some connecting planning decisions need to be made between the phase and thread levels, as suggested above. If an individual thread involves a sequence of sessions viewed from the perspective of a common element (investigator, informant, etc.) then we can also see each phase of the KA effort as having the task of scheduling the order in which these various elements are deployed or studied.

With the various threads of the canvas in hand, we can proceed to consider the planning of an individual session as it sits at the intersection of its various threads.

4.9 Planning a Session

Within the overall context of a KA enterprise, a number of individual KA sessions will be conducted. Although global decisions will be made as part of broader KA enterprise- and thread-level planning, many detailed planning decisions need to be made on a session-by-session basis. This section will present a framework for session planning which addresses what are probably the most complex kinds of sessions from a planning perspective: interactions involving groups of multiple informants facilitated by teams of investigators. Planning requirements for other kinds of sessions such as one-on-one interviews or artifact analysis may allow for a simpler set of variables to be accounted for.

An overall "architecture" to a session can be defined, which involves the following steps:

1) Initial planning and scheduling of the session, including: specifying session objectives, selection of topics to be explored, participants, desired format for resulting workproducts, etc.

2) Preparatory work for the various participants.

3) The primary knowledge elicitation event, i.e., interview, meeting, observation session, etc.

4) The "write-up" or codification process, usually involving some interpretation time on the part of investigators after the immediate event;

5) Validation of documented results as required with the informants that participated in the KA activity itself, and possibly with other members of the focus community. (Note that this can spill over into another session context.)

6) Dissemination of the codified and validated results, via the dossier, to other participants in the KA enterprise.

7) Folding results of the session back into KA plan to update the various threads and adjust ongoing elements of the plan.

Planning aspects of these various elements of the KA session are discussed in the paragraphs below.

4.9.1 Establishing Session Objectives

Planning a KA session first and foremost implies making decisions about the basic elements of a session, including the following:

- Objectives/Topics of focus: What is the primary purpose of the session? What topics are of interest? How will this session build on, and advance, the knowledge gathered for those topics in preceding sessions of the KA enterprise?
- Audience: Who will be the primary audience for the resulting codification of session results? In particular, what validation from the focus community will be required? What sharing of interim results among the investigator team will be necessary?
- The setting(s) of interest: Is a particular setting going to be the focus of the session? Might there be parallel sessions planned to elicit similar descriptive data about different settings? Or is the setting being used to provide a "normative" view of domain-specific practice?
- The investigator(s) conducting the session: Who is available, and best qualified, to conduct a particular session? How will that session develop the threads for the various investigators involved? Is the intention to provide them focused experience in a few topics or settings, or to expose them to a diversity of perspectives?
- Informants that might be involved in the session: Who is available, and who best suited to provide information on the topics of interest?
- Inputs: What artifacts or previously developed KA workproducts might provide prior background information for the session? Which could be used directly as knowledge sources and/or a basis for interactions during the session itself?
- Results: What are the anticipated KA workproducts to be produced from the session? What formats or representations should be employed?

By considering the various links from the session being planned to the thread level of the overall KA plan, the session planner gains an understanding of the motivation for the session from the standpoint of each intersecting thread, and the impact of the session as enacted upon each of those threads (and any threads newly created, such as the threads for those workproducts created as session results).

Clarification of these objectives aids in making one of the fundamental session planning decisions: what format of session is most appropriate given the participants and objectives? Some options are discussed below.

4.9.2 Selecting Session Format

So far, we have treated sessions as if they were controlled interactions between a single investigator and a single knowledge source. In fact, a single session could manage an interaction among several elements in the KA effort. Many of these types of interactions are familiar, either from everyday work practice, or well-known knowledge and requirements engineering practice. In fact, a healthy mix of different session types is likely to result in a more flexible KA effort.

The following paragraphs outline some possibilities of different types of KA sessions. This is by no means an exhaustive nor systematic list. The main point is to see how the basic Canvas framework (the interaction of elements) applies in each case, and to appreciate how a healthy mix of different session types can improve the overall KA effort, by suiting the session type to the kind of knowledge desired and the needs and predilections of the various participants.

One on one interviews

The individual interview, with a single informant and a single investigator, is probably the most commonly used session configuration in expert systems development. When the informant has highly specialized expertise, it can be the case that there is only one informant available, and a one-on-one session is a necessity. It has the advantage that the session can focus on the point of view of a single informant, without the risk of moving into discussions between experts. When the informant is accustomed to a teaching or instructional situation (as if often the case with recognized experts), the one-on-one interview can easily degenerate into a personal delivery of material that has already been prepared, thus failing to deliver the embedded knowledge that is aim of a knowledge acquisition effort.

Facilitated group KA sessions

Group KA sessions have a number of advantages, not least of which is that some cultural interaction in the group can take place with the investigator present, allowing him the possibility of direct access to culturally embedded knowledge. Facilitating such a session can be challenging, since the informants have already established some relationships, terminology or habits that might hinder the knowledge elicitation process. For example, informants whose status in the work setting is lower than other might not feel that they can express themselves in a group session. The session might move into a technical discussion of details that are out of scope of the knowledge acquisition effort, using terminology with which the investigators are not familiar. Despite these and other dangers, group sessions are often the mainstay of a large knowledge acquisition effort, because of the economy of effort they provide. Group KA sessions often involve partnering on the part of investigators, which can help to offset the gaps in knowledge of individual investigators.

Walkthroughs

An investigator studying an artifact alone will often have several questions about the context in which it is used, the meaning of the terms, how it relates to other artifacts, etc. One way to handle such a situation effectively is to plan a session in which a human informant works with the investigator as he studies the artifact, and "walks him through" the various pieces. This is particularly useful when the artifact has a process nature, such as a procedures guideline manual.

Demonstrations

Closely related to the walk-through, especially given, as described in Section 3.3.2, that knowledge acquisition will often occur in technology intensive settings, is the program demonstration. In this case, the artifact to be studied is a piece of software, which is run during the demo, and is

accompanied by some explanations from an informant. This format provides a great deal of possibilities: multiple investigators (as in a demo performed for a group), specific questions and answers, and limited experimentation on the part of the investigator. Experimentation can be either planned or not. For example, during a demo of a system that diagnoses engine problems in an automobile, the person giving the demo might ask the audience, "Think of a problem with your car." Or, if the investigator is in a position to drive the demo, she might ask, "What would happen, if the program were to receive this input?". Demonstrations can place practitioners into a context with which they are familiar, thus facilitating knowledge acquisition.

Automated sessions

An extreme example of a knowledge acquisition session is the fully automated session, in which a human informant is left alone with a computer program that leads him through the session. Manusco and Shaw [23] report that an automated session is often liberating for an informant, because the informant has a better feel of ownership over the resulting knowledge. She also reports an increased sense of trust, that the machine will not judge the completeness or coherence of the knowledge. Weizenbaum [57] reports the uncanny success enjoyed by the Eliza program in gaining the trust of its users, even to the point that they would ask other people to leave the room while using it. Although this is a promising approach, the current state of the art in fully automated knowledge acquisition is such that only a limited set of highly specialized representations can be produced in this manner. Other automated sessions include automatic processing of documents, such as word frequency counts, which provide objective and sometimes surprisingly insightful information about natural language documents.

Wizard of Oz sessions

A common method used in human computer interaction studies, expert system design, and requirements engineering is the so-called "Wizard of Oz" method, in which the investigator pretends to be a computer system that will be introduced into a workplace. The informant deals with the expert as he would with the system. The behavior of the system can be changed simply through agreement between the investigator and the informant. Although the method has been traditionally used to project the impact of new technology on the workplace, it can also be used to uncover the hidden assumptions in the work practice.

Observation sessions

A very useful, though sometimes demanding method for acquiring knowledge of a work setting is through observation, that is, the investigator is present at the site of the work practice and observes it in action. One possible problems stem from the fact that practice is likely to change under outside observation. The solutions to this problem include having the investigator enter the work setting in some accepted role, often as an apprentice or junior colleague, and make observations while actually performing in the work practice in the focus community. Another solution is to provide automatic surveillance, through video or audio recording machinery. In technology dependent settings such as those described in [3.3.2], the technology itself can provide opportunities for automatic observation.

4.9.3 Session Preparation

Once the objectives, participants and format for the session are established, it is important to consider what advance preparation may be required for each respective participant in order to achieve the session results. This includes preparation of the informant, the investigator, and the setting, among other elements.

Informant preparation

Where is the informant in his thread? Has the informant received project orientation/familiarization? Has s/he participated in previous KA sessions? What additional information does the informant need to know beforehand?

Will s/he be asked detailed questions about an area in which he may have spotty recollection? Will he want to do some review work before the session? Or is this something specifically to be avoided? One KA objective (or guidelines) may be to ascertain "knowledge in use" rather than the official knowledge formally available.

> Example. In interviewing a helicopter transport pilot for TCIMS, the interviewers asked about an equipment maintenance checklist that hangs on the wall in the office and is used to monitor scheduled maintenance for the helicopter. The informant mentioned some example categories of maintenance tasks from memory in describing the sheet. Different data would have been obtained if this topic of interest had been previously called out in the interview preparatory material. The informant might have reviewed the material to give a more comprehensive description, might have been more cautious if there are legal or contractual restrictions on "correct practice" that are sometimes relaxed in practice, or might have simply told the investigators to go look at the document itself.
>
> In this situation, the un-prepped responses yield possible data about "foreground" versus "background" maintenance problems, or "typical" versus "atypical" problems. Naturally, there is a danger of over-interpreting. Choices made by the informant about which maintenance problems to mention might be related to relative frequency of occurrence, relative criticality of the problem, the length required for a fix with its concomitant impact on equipment availability, the informant's expectation of what examples would be understandable to a non-mechanic, or any number of other factors.
>
> Thus validation of some sort would be necessary. This could be done within the same interview, in a follow-up interview, in a request for comment included with a transcript submitted to the informant for review, or by cross-checking in interviews with other informants. The interview results could also be cross-checked against the artifact itself. In this case it would be of particular interest to view the artifact in context to see how it is positioned, what marginal notations are made, etc. Theoretically results could also be checked with a maintenance mechanic. However, here the question of domain focus becomes relevant. Since the overall objectives of the TCIMS project are focused on emergency trauma care, the transport domain is itself only tangentially related; the domain of helicopter equipment maintenance would be an even more remote "second cousin" and probably out of scope for the KA effort.

It might be desirable to let the informant know what kinds of questions will be asked beforehand, or to allow him to prepare some answers to general questions, to get the session started. Or this might be intentionally avoided, so as to avoid having the informant look up official answers, rather than giving his own.

> Example. In the TCIMS interview of the helicopter pilot, a template was given to the pilot before the interview, describing the nature of the material to be collection (task information), and asking some general questions about major responsibilities, durations, etc. Any answers given on this form were used as starting points in the discussion. In the case where no answers were given, the interview would begin with similar questions from the investigator.

Investigator preparation

Where is the investigator in her thread? What information pertaining to the session topic could be acquired from basic sources, including well-known texts or other KA session reports? What information could bias the investigator inappropriately for the session?

> Example. In the TCIMS interview of a helicopter pilot, the investigator brought with her knowledge of the structure of the emergency room and hospital for which the pilot worked; in particular, the administrative structure of the transport service (as a private contractor, hired as a company by the hospital) had considerable impact on understanding the impact of rules and regulations governing the work. This information had been gained earlier in the investigator's thread by her participation in other interviews with personnel from the emergency room transport service.
>
> In contrast, it would not have been advantageous for the investigator to have begun the interview with detailed knowledge of helicopter flight, since the topic of interest in the interview was the responsibilities of a medical transport pilot. Knowledge of helicopters could have biased the discussion toward details of the mechanics of flying, rather than the interactions with the medical constraints of the situation.

As a general rule, before holding a session read/skim materials relevant to the session; in particular, any materials written by or relating to the informant(s) involved. Even if the informant doesn't expect you to have read the materials, terminology may be used that is explained by the material. If possible, prepare a few detailed questions that show attention has been paid to the materials, but don't try to misrepresent the preparation you've actually done.

Some preparation of the setting itself will also be required: e.g., arranging facilities, notifying other people that the session will be taking place, etc. This preparation is particularly important for observation.

4.9.4 Session Performance

Full exploration of the many subtle elements involved in conducting interviews, joint design sessions or artifact analysis are beyond the scope of this guidebook, which is oriented primarily towards the planning and dossier management aspects of KA. This sub-section highlights a few aspects of KA session performance that very directly reflect key Canvas principles at work.

Rapport

Building rapport is an essential part of any productive human interaction, but in KA it is critical for developing the trust needed to validate information and enable mutual knowledge building and discovery. If relationship is the end and trust is the vehicle, then rapport is the starting point. Planning and conducting sessions so that rapport can happen and is given appropriate emphasis at different points during the session is an often-missed and undervalued aspect of KA planning.

The beginning of the session is most critical, perhaps the first 20 minutes. In these first minutes, the investigator will benefit from damping his curiosity about facts and content, while focusing on the person's background, the physical environment, and nonverbal communication cues. For example, this usually means emphasizing eye contact over note-taking. In short, it means making a connection first and foremost.

One way of thinking of a KA session is illustrated in Exhibit 10. The picture shows that early in the session attention to exchange of information is low; development of rapport is high. They

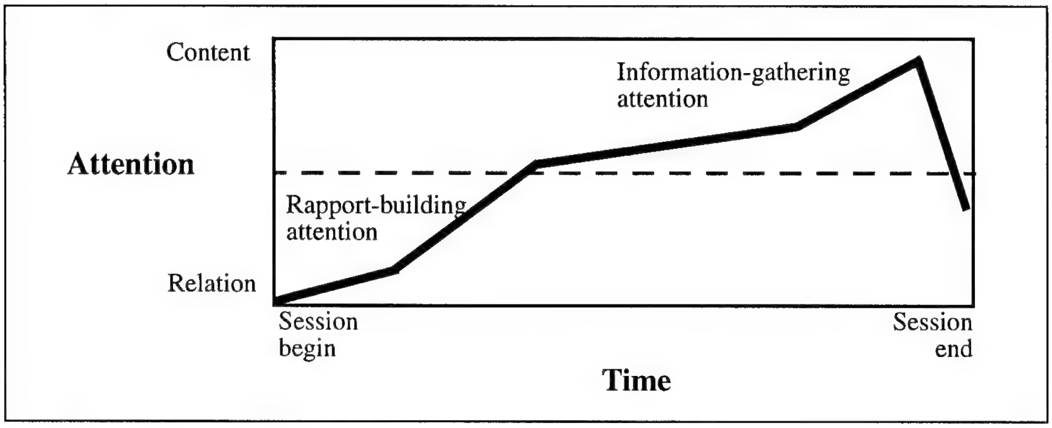


Exhibit 10. Attention in a Session Devoted to Rapport Building versus Information Gathering

become equal roughly half way through, and shift toward the end. When sessions are going well, they equalize about one-third of the way through, shift toward information exchange about two-thirds of the way through, and the final third opens up the possibility of joint discovery, knowledge creation, and possible knowledge modeling. At the very end of the session, there should be a re-framing of the relationship to avoid an abrupt, task-oriented cut-off to the session.

Thinking in terms of these proportions can be significant session design considerations. For example, some researchers may need to plan a set of more personal questions for the opening. Key or controversial questions should be saved until rapport is clearly being established. Pictures or visual models should not be introduced too soon, breaking eye contact and connection between informant and researcher. Finally, ending with plenty of time for creative exchange to happen spontaneously should be planned. Sometimes it is just as the session is ending that an informant feels greatest trust and wants to give most. (Note that the exhibit is *not* a graph of energy level.) These are just examples of how session planning can be enhanced by considering the structure of rapport building.

Session logistics

A number of logistical problems can arise in planning a session, including gaining access to the informant at all (in military settings, it is important to observe the chain of command to assure that access to the informant will not be denied at the last minute), planning a location for the meeting, arranging for any equipment (recording devices, demonstration machines), scheduling, etc. It is worthwhile to consider why logistics decisions were made as they were, so that if something goes wrong (e.g., a technical problem with a video recorder), a timely decision can be made about how to react.

Context setting

The informant must be oriented as to the nature of the KA enterprise and the larger project of which it is a part. This could be done as part of the preparatory activity for the session, but it is a good idea not to make assumptions that any preparatory material has actually been reviewed thoroughly. For example, for one interview which we observed, the material sent in advance for the informant's benefit had never been passed on. This might be typical of expert informants in high-pressure, time-constrained environments.

Open versus closed session facilitation

There are trade-offs involved in the session planning process. For example, insisting on a strict focus for an interview may help keep the interview on track, or may blind the investigators to essential information volunteered by informants which happens to be out of scope of the session objectives. The session objectives should help clarify this. If the purpose of the session is baseline familiarization for the investigator, or getting an informant "on board" as an interested contributor to the project, more flexibility might be appropriate in selecting an open versus closed style for the interaction.

4.9.5 Session Follow-Up

The following paragraphs discuss major follow-up activities subsequent to the session event. Depending on the extent to which collaborative modeling was done in the session, some of these activities (particularly write-up and validation) may have been incorporated into the session itself.

Writing it up

Perhaps the most difficult part of a knowledge acquisition session is the write-up; this is the work-product that will be entered in the dossier and serve for some audience community to know what happened during the session. One of the major problems is to ensure that as much of the information that was elicited during the session as possible makes it into the dossier; this can be simplified if the session itself followed strict goals. However, as mentioned above under Open versus closed sessions, there are other reasons for the session to be less structured.

In writing up the session, it is important to keep in mind the audience for the write-up. If the write-up includes any representations, these must be expressed in a manner that is accessible to this audience. Representations were identified with their target audiences in Section 4.6; this information can be used as a guide in selecting representations to use in write-ups. There might be need for several write-ups, for different audiences, in particular, one for the investigator team, to help them gain understanding of the dossier, one for the informant, to verify the information, and one for the target audience.

Validating a session

Interviews with informants should be validated in a special session with that informant. Validation of artifact studies might involve consulting an informant, though sometimes a review by the investigating team might be sufficient. This could be as simple as having the informant review the write-up, or as complex as having a follow-up session to examine the details of the critique. The trade-offs here involve the informant's time and commitment to the project.

A particular difficulty arises in validating representations of variability, not only because of unfamiliar representations, but because such models do not represent the "world view" of any one informant who could validate them. Special validation techniques may be required. For example, multiple informants from different settings could be brought together to validate a model that captured the range of variability in the knowledge gathered from those settings. Applying such techniques may have the side-effect of changing the way informants organize and reflect on their own domain knowledge. In this case, the KA and modeling process is a definite intervention in the dynamics of the work setting, whether or not the final result is introduction of new technology into that setting.

Dissemination of session results

In the Canvas framework, the primary means of dissemination of session results in through inclusion in the dossier. This includes establishing linkages to all new artifacts accessed as part of the session, and placing any generated new workproducts (session notes, formal models produced, etc.) in with the appropriate indices.

The information may also be directly presented or transferred to appropriate recipients in the target community (e.g., technologists viewing interview reports). Each phase of the knowledge acquisition enterprise should conclude with some definite hand-off of this kind to an audience in one or more of the target communities. However, in most projects the hand-off from KA is to a more extensive modeling process, where data will be integrated, synthesized and formalized in a variety of ways.

Updating the Knowledge Acquisition Plan

As the final stage in the life cycle for each session, the new learning from the session is used as a basis for updating and possibly adapting the KA plan. In particular, all the threads for each participant in the session need to be updated. Artifacts that were studied now have an additional pass of review and interpretation. Informants will have undergone an additional KA session; they will be more familiar with the process, and perhaps ready for a more intensive session; or perhaps the allotted budget for their time has now been consumed. Investigators will also have progressed in their respective threads.

There may also be new elements that will have new threads established for them in the plan. For example, any workproducts generated at the session may be beginning a series of transformations into more suitable representations for the target audience; their threads are initiated here.

Finally, after updating all the various threads, actual re-planning activity can take place. How this is coordinated in terms of the multiple sessions that may be occurring simultaneously, with a large-scale KA project, is a matter of project management. At some point, though, the various changes in the elements tracked in the KA plan, some anticipated, some not, need to be addressed by re-visiting and possibly modifying the plan. As a simple example, a given informant may suddenly have revealed areas of knowledge that were not accounted for previously; now new sessions might be scheduled with this person. Or new leads for other knowledge sources will have been obtained. In addition to such expansion of the plan, there will be dead ends and *cul de sacs* reached that should be trimmed from the plan as soon as possible.

4.9.6 Issues in Session Planning

A plan for a knowledge acquisition session has to deal with a large number of competing constraints and trade-offs. These issues can affect any aspect of the session plan, including the structure of the session, the participants, and the representation used for the write-up.

Dynamics of single versus multiple informant interactions

Consider the decision to interview a domain expert in a one-on-one setting versus within a group. Canvas provides guidelines for identifying some of the key issues that should be considered. Social interactions will occur between multiple informants. This could result in additional information emerging from the interaction which could be of value to the KA enterprise. If the multiple informants are co-workers from the same setting, anecdotes and recollections of past experience may come up as part of the interaction that an external investigator would have lacked

the context to introduce. These particulars might yield more general principles relevant to the topic of focus. This reflects the principle that important types of knowledge are carried in the social interactions that make up the dynamic aspects of any community of practice.

However, communities have defined social relations, and these relations will affect what kinds of knowledge emerge from the interactions and what aspects of this knowledge people will be willing to share. If there are line-of-authority relations among the informant group those in a lower staff position might be unwilling to voice opinions that would be deemed critical of the management, whereas in a one-on-one interview they may speak more freely, especially if guaranteed confidentiality or anonymity of attribution. Conversely, those in positions of more authority might be loathe to express uncertainty about future business in an exposed setting.

From the thread perspective, an investigator's level of experience and degree of credibility within the focus community (established over a series of interactions) would need to be considered. A multi-informant knowledge elicitation session runs the risk of turning into an interaction between the informants, and the investigator might need to be prepared to step into the role of mediator or facilitator to some extent. From an informant's thread standpoint, it might generally be a good idea to interview an individual first, in order to predict how his or her interaction styles might affect others in a KA setting. On the other hand, a group interview with very clearly stated goals and structure could be used as a screening device to earmark promising candidate informants for further work.

Evaluation of investigator results

Work products created by investigators may be subjected to multiple types of review. One of the most valuable is review by the informants themselves, with attention to accuracy and completeness. Further review may be carried out by peer investigators, with consideration to clarity, attention to established investigator processes, suggestions for follow-on KA activities. Another important form of review is by an established group of experts as in TCIMS that takes a more global look at the body of KA results and the credibility of the sources of the information. Finally, there may be reviews by technologists with respect to the usability and formality of the information, its suitability as source material for work products such as software requirements and design documentation.

KA enterprise input to session planning

Many of the decisions made while planning the enterprise can be used as guidelines in planning the individual session. We will demonstrate this with an example from TCIMS, where stakeholder issues, the planned audience groups and their associated representations (see Section 4.6) all have an impact on the successful planning of the sessions.

> Example. Two lessons learned from the TCIMS project concerned the accessibility of information to various target groups. On the one hand, the group of technologists (i.e., the system developers who will be introducing new technologies into the field) were reluctant to recognize either the difficulty of the knowledge acquisition effort, or, more importantly, its value to their task. Technologists complained that there was too much verbiage in the material, and that it did not tell them what they needed to do to develop their systems. On the other hand, representations used in some KA products were unfamiliar and perhaps even alien to the medical community. The effect of this was to increase the difficulty of model validation.
>
> Both of these types of problems can be ameliorated, or at least better anticipated, by paying close attention to the intended audience of each workproduct. This would both reduce the amount of information that the technologists would have to deal with, as well as make sure that the information that they do access is understandable to them. Similarly this would

remove the necessity for members of the medical community to learn representations that are unfamiliar to them.

The lesson to be learned for the knowledge acquisition project planner is that it is necessary to explicitly know the intended audience of each KA workproduct, and to tailor each workproduct for its intended audience. In particular, a KA workproduct that will be accessed by multiple audiences may need to be maintained in multiple forms. Here, automated tools to support seamless transfer of information among representations would clearly be a boon, as the collaborative nature of the KA enterprise means that changes can ripple back and forth along the life cycle from raw data acquisition to more formal modeling and interpretation of the data.

Variability issues

In general, in dealing with a given informant, we can distinguish between variability that can be directly elicited from his experience and variability that can be modeled only by aggregating or integrating his experience with other data. More specifically, we can distinguish the following cases:

- In some cases, particularly reflective or expert informants can offer data already close in form to a model of commonality and variability for their domain: an abstraction of their experience. In many stable domains, elaborate codifications of data are available (e.g., in the medical domain, established protocol or procedures documentation). Such data needs to be handled in the KA process as an artifact, but as a particular kind of artifact which provides second-order, interpretive or meta-data as it occurs within the focus community.

- Variability that is part of regular work practice for the informant but is articulated through or as a result of the KA interaction. This could be documented as part of a work process model, procedures (with contingency conditions,) etc. These could range from truly routine, low-skill activities to activities requiring expert judgment. For a practitioner such as a medical staff person, triage procedures are an example. Here, the knowledge is codified as a result of the KA process but was known or accessible to the informant prior to KA participation.

- Variability can be elicited from an informant's experience through a process of reflection. The informant offers observations of variability on the basis of varied experience (moving jobs, working at different sites, etc.). The KA interaction may trigger a reflection or learning process that changes the informant's relationship with their own knowledge or practice.

- Variability observations result from making new information available to the informant and thus sparking reflection that is a direct intervention of the KA process. This could include walking an informant through a set of artifacts similar (but not identical) to those with which he is familiar; bringing together a group of informants filling similar job roles to compare notes; or asking an informant to validate a KA workproduct which represents codified variant data gleaned from other informants. As with the case above, the learning involved in sparked by the KA interaction but in this case more of an intervention has taken place, since new data is made available to the informant.

- Variability is observed by the investigator in synthesizing data obtained from multiple informants, artifacts and/or settings (often the assumed scenario in KA).

  Example. Two medics might have different ways of classifying the various triage decisions that they face. In this case, we must compare not only a set of triage instances or cases which we then synthesize into our own classification or comparative model, but also the way the two medics have differently "made meaning" out of the range of cases in their respective bases of experience.

The examples above range over several different factors. The primary factor is the extent to which the knowledge is already extant in the focus community, is created via interpretation by the investigators, or emerges through collaboration between informant and investigator. The KA plan should make necessary distinctions between these different cases and track representations differently depending on their origin. Otherwise it is difficult to correctly interpret the representations or models.

Each such model, whether a previously created informant model treated as artifact, or a KA workproduct created in collaboration with investigators, embeds a "theory" about variability in the domain. There may be multiple such models, or a single such model may diverge from the theory that emerges from the investigators' own models in the same area. This becomes a model resolution issue that reflects variability at a meta-level from that observed in the work practice of the domain itself. Some social scientific academic work has been done in relevant areas such as meta-ethnography that deal with the synthesis of multiple, qualitative interpretive data analyses [28]. In general, though, this is an area requiring substantial further research.

Representations for representing knowledge acquisition need mechanisms for describing variability. These will be discussed in more detail in Section 5.0, which discusses representation strategies for KA.

Investigator bias

The investigator's previous experience and degree of relevant expertise can be a significant factor to consider in planning. Bias on the part of the investigator can affect the accuracy, consistency and completeness of elicited knowledge. The effects of bias can be anticipated during enterprise planning, and the development of the bias itself can be tracked during thread planning. The effects listed below can be counteracted to some extent during session planning:

- Missing data (did not ask the question).
- Inaccurate data (mis-heard, mis-reported, imposed own view).
- Interaction led to false data (led the witness, created resistance, or created an over-willingness to please and give the "right" answer).
- Mixed source of data (uses session as vehicle to interject own opinions and knowledge without acknowledging this).

Strategies to counteract risks of investigator bias include the following:

- Having informants interviewed by teams that include an investigator with a high degree of topic expertise and a relative novice (with respect to the topic).
- In Section 4.5.3, the investigators were characterized according to their skill levels in specific techniques and procedures for suspending unwanted influence of their expertise on a session. Selecting investigators with these skills for sessions where the effects of bias are particularly risky or likely.
- Build into the session structure explicit solicitation of feedback from informants about their perception of how well the investigator has performed their task. Were their remarks properly understood and noted? Were important points missed or glossed over?

Informant bias

Bias on the part of the investigator can skew results in various ways, as described above. Bias on the part of the informant is slightly different in nature, as outlined in [26]. Of course, an informant will be presenting personal opinions and viewpoints; this is a given in dealing with the knowledge source. The problems here can take the following forms:

- Lack of clarity in distinguishing personal knowledge and opinion from consensus knowledge of the community of practice.
- The "say versus do" problem: what is reported may not correspond to actual practice.
- Unwillingness to share knowledge, or distortions ranging from "white lies" and politically correct answers to deliberate falsification.
- The informant may have a lot invested in the uniqueness of their knowledge and problems; or conversely may be too ready to report in generalities and "normalize" practice.

Different strategies for planning sessions can adjust for some of these factors. Relevant factors in selecting the right strategy include the closeness of the session's context to the work setting about which knowledge is required, and the degree of awareness of the informant that knowledge acquisition is occurring. These various strategies can be viewed as a kind of spectrum of options:

- The most typical interaction between an investigator and an informant would be in a formal interview outside the work setting (e.g., in a conference room, perhaps with a tape recorder or video, certainly with the investigator taking notes during the session). If the informant is being asked to recall detailed procedural knowledge this may be an awkward setting. (The scenario elicitation process is one attempt to facilitate recall by appealing to natural ways of conceptualizing work flow and task boundaries.)
- An interview conducted in or near the work setting itself may improve the closeness of fit between the reported and actual practice. Passive observation of work actually being performed, with the investigator visible and acknowledged but otherwise not directly interacting, would generally yield richer data along these lines. However, the activities in the work setting will be affected by the presence of an observer.
- Participant observation, where the investigator plays a work-related role within the setting of focus, can be less obtrusive because there is an acceptable role for the investigator within the normal roster of work-related roles. This personal, experiential data can result in far richer knowledge transfer. How much of this is effectively *codified* depends on the skill of the investigator.
- Non-intrusive data collection (e.g., instrumented tools for data collection and logging) would in theory enable obtaining more accurate information about real versus official practice. Here an ethical issue arises. In theory, clandestine observation would yield the most objective data, yet it would generally be unethical to collect and utilize data on this basis as part of a knowledge acquisition effort. When informants are aware and willing to participate, non-intrusive data collection has the advantage of not "breaking the flow" of the event.

4.10 KA Capability Development Plan

We have stated that only in a few cases will the investigators in a KA enterprise have true stakeholder interest in codifying their own skills in KA for transfer to other domains. Nevertheless, for any project of significant size there are good reasons to include a systematic means of assessing and improving fundamental KA skills among investigators. For any informants who have exposure to multiple investigators over time, presenting a reasonably consistent standard of approach and competence is important for the overall credibility of the project.

In this section, therefore, we offer some suggestions for questions that can be used for self-assessment and as a basis for ongoing evaluation, exercises and discussion among investigators.

- *Do I know how to "play dumb"?*

  When I am interacting around a topic I do know well, how good am I at suspending my own knowledge and opinions, in order to elicit the knowledge of the informant? (Imposer vs. good facilitator?)

  In order to allow an informant to offer information in a form with which he is comfortable, it might be necessary for an investigator to allow the informant to report information with which the investigator is already familiar. By acting as though he does not already know the things the informant is reporting, the investigator can encourage the informant to cover basic material thoroughly. This could well allow the session to uncover some normally unspoken assumptions that are embedded within the work practice.

- *Do I know how to "play smart"?*

  When an informant moves into territory that is new to the investigator, there is a good chance that a lot of knowledge can be acquired. If, however, the investigator seems to be getting lost, this will encourage the informant to retreat to firmer, simpler ground. An investigator who can give the impression of understanding all the intricacies of an explanation can encourage the informant to move into new areas. The investigator should be able to:

  — Ask interesting questions when still in the dark

  — Tolerate not knowing

  — Maintain informants' buy-in when they may tend to be intolerant of people not familiar with their field

  — Project an air of intelligence in areas where their knowledge is insufficient

  — Know when and when not to try to bluff

- *How well can I skim/filter technical material?*

  While not often discussed as a skill, one of the hardest tasks that faces the investigator in KA is how to rapidly filter information sources. The temptation, of course, is to get lost (and hence, for any non-trivial domain, overwhelmed) in the detail; or, literally, to not know how to extract useful points from the morass of material. This is a learnable and practicable skill, that requires attention to both time- and topic-based constraints.

- *How attentive am I to reflecting back the terminology of informants in the interview?*

  Do I introduce my own terms and expect them to shift? Do I check out and verify whether my usage of unfamiliar terminology is accurate?

- *How fast am I at picking up knowledge in new areas? (Quick study?)*

There is fascinating interplay between these acquired skills (or traits) in knowledge acquisition. For example, does a person who can readily suspend their expertise in a familiar area also tend to be a person who can facilitate well in an unfamiliar area? We speculate that good listening skills are a fundamental requirement in both cases.

We posit a knowledge acquisition skills "maturity model" that illustrates the various levels of skill that an investigator might already have reached by virtue of other, similar activities:

- The first "level" of knowledge acquisition skill recognizes that many of the skills of a good reporter are relevant to knowledge acquisition, such as building rapport, remembering what is said, taking effective notes, balancing leading and following in the interview, and taking care to cross-check and verify stories. Good reporters also have less tangible skills like "smelling a good story," being resourceful about finding people to talk to and other sources of information, etc. Ethical questions that arise for reporters such as balancing the impact of a story on the subjects and their community with the public's "right to know" can be understood in terms of conflicting stakeholder interests. Investigative journalism and detective work involve many of the same skills. Note that in classic journalism there is no imperative to report using the terms and concepts of the focus community; the reporter acts as the "eyes and ears" of the larger culture.

- The second "level" of knowledge acquisition skill can be differentiated by going beyond news reporting or detective work in that it attempts to elicit knowledge from information sources (informants) without imposing concepts, categories, or language from the reporters' or detectives' culture. This kind of concern would be more typical of a professional ethnographer (e.g., a cultural studies researcher) for whom capturing the native or emic categories is a specific goal.

- Unfortunately for the goals of knowledge acquisition, each session is inevitably an intervention in the focus community. The next "skill level" involves not merely seeking to have no influence, but acknowledging and attempting to plan and compensate for the influence that occurs. An inter-cultural reflectiveness is required to recognize what categories, values, and language the investigator brings as a member of the community doing the research (in addition to her personal biases). This is particularly important because it not only colors the data acquired, but it does so in a way that is often very difficult to see in the data itself. To become aware of these factors means owning up to the politics of how "we" who are doing the studying may impose on "they" being studied. Sometimes this type of research reveals as much about "us" as it does about "them." In Canvas, the emphasis on looking at all the various stakeholder issues involved is one step towards fostering this skill level in KA.

- In what we view as the highest skill level in KA (within the scope of this document) the informant and investigator are engaged in true collaborative knowledge acquisition, i.e., they see themselves as working together to create a codified model of knowledge in the topic areas. True collaborative knowledge acquisition requires that investigators and informants engage in a relationship that (1) temporarily suspends the cultural lenses of their respective work settings; and (2) builds enough trust to discuss topics that may be culturally "taboo," politically sensitive or so deeply internalized as to be hidden to practitioners.

Skills required for collaborative knowledge acquisition

Collaborative knowledge acquisition expertise includes the skill of being a "skillful non-expert" who can help experts reflect on and articulate their own knowledge in fruitful ways. The investigator many times must act more as facilitator than an interrogator, by knowing the right questions to ask. The investigator must also, in the terms of Donald Schon, be a "reflective practitioner" [31] continually practicing the intensive inner work of examining and adjusting for personal and cultural bias. She must be ready to experience the ground shifting in the interview, discover and suspend successively finer details of this bias, without feeling defensive or, alternatively, fleeing her own cultural bias in favor of the bias of the informants. (This last point highlights the risk of loss of perspective that could be called "going native.")

The investigator must hold her own need for closure in suspension in the face of emerging data that is not immediately consistent, clear, or coherent. Maintaining high receptivity without judgment or premature conclusion in these situations is a difficult and a critical skill. The investigator builds a trusting relationship, but in a way that is not exactly mutual and does not allow the "interview" to lapse into normal friendly conversation, problem solving, advocating, or advising.

When the informant temporarily leaves the flow of his cultural daily life to reflect, talk and allow his knowledge to be elicited by the interviewer (who is suspending the lenses of his own culture described above), then a third, "collaborative" vantage point is established. This could be thought of as a temporary culture in which both informant and investigator can "play" as equals, confronting even harsh realities and envisioning new possibilities. This is where innovation may take place in a way visible to and acceptable to both communities of practice.

The description above places demands on someone who performs knowledge acquisition that may take years to master. From this set of questions, one might think no human ever gained the objectivity, creativity and receptiveness necessary to be an investigator in a knowledge acquisition project. Although we have presented this in terms of skill levels, the analogy to a true maturity model may be misleading here. Different contexts impose different needs for handling various issues (e.g. a reporter need not aspire to be an investigator in a KA effort.) It is possible to carry out knowledge acquisition tasks without having mastered all the subtleties of a collaborative, ethnographic approach. Experience with other related activities can develop some of these skills, and investigators can be effective even as they develop and refine the skills outlined above.

5.0 Representation of Knowledge

Representation of knowledge forms the cornerstone to knowledge acquisition. Since knowledge acquisition refers to transfer from one community of practice to another (rather than to a single individual), the knowledge must eventually be written down in a form accessible to the target community. One of the major skills of an investigator is fluency in systems for writing down knowledge. We refer to such systems simply as representations.

A ***representation*** is very similar to a language. Not only is it a system for writing things down, but also, like a language, it includes two components: a *notation*, in which the representation is written, and a *semantics*, which provides meaning for workproducts written in that notation. However, in contrast to most languages, representations for knowledge acquisition typically have a much narrower scope of expression. Typically, the notation of a representation provides a small number of elements from which to construct workproducts; for example, boxes and arrows, or arrays of cells. The semantics of a representation specifies how these elements are mapped into some real-world things, or relationships between real-world things. As we shall see, it is this narrowness of scope that allows representations for knowledge acquisition to provide guidance for the knowledge elicitation process.

Representations play several roles in the knowledge acquisition process:

- Within a single community of practice, representations can be used as a part of everyday work, or as part of the training process for introducing new members into the community. This applies to both the focus and target communities in a knowledge transfer configuration. An investigator must have the flexibility to be able to manage workproducts expressed in these native representations.
- As part of the knowledge acquisition process, the investigator will produce workproducts, which will have as their intended audience members of the focus community, e.g., for purposes of validation. Such workproducts will have to be produced in a representation accessible to that community.
- During a particular knowledge acquisition session, a representation can be used to focus collaborative work between the investigator and one or more informants. Using a particular representation during the session keeps the focus of the session clear, and allows the investigator to move systematically through the material.
- Also as part of the knowledge acquisition process, workproducts will be produced with the target community as intended audience. These workproducts will have to be expressed in a representation that is accessible to the target community.

If we think of knowledge acquisition as a craft, then representations are the craftsman's tools, and the mark of a skilled craftsman is fluency with these tools. No single tool can cover the entire range of jobs; just as a craftsman has a kit of different tools, an investigator should be familiar with several representations. In the Canvas context, a ***repertoire*** of representations is a set of representations used by a specific community of practice. The repertoire includes both the representations and the relationships that enable their use together or in complementary ways.

In order to have command over a repertoire of representations, an investigator should know the following things about each representation:

1) How to create and maintain workproducts written using the representation,

2) What circumstances indicate and counter-indicate the use of the representation,

3) Common traps that users of the representation can fall into,

4) The built-in bias of the representation, and

5) How the representation is related to other representations in the repertoire.

The idea of a repertoire of representations cannot be stressed too highly here. Skilled knowledge engineers[1] should develop a broad repertoire of representations to allow them to interact with many different practitioner communities. We have seen numerous knowledge acquisition projects run into trouble because a knowledge engineer has studied up on a particular representation, and has insisted on continuing to use it, even after it was apparent that nothing was to be gained from that representation. A good craftsman can recognize the right tool for the job.

Just as a craftsman, throughout his career, acquires a collection of tools that he feels comfortable using, and with which he is most effective, a knowledge engineer also collects representations with which he is comfortable. It is not our intention in this book, nor would it be possible, to offer a definitive set of representations that we expect all knowledge engineers to use. Instead, in the rest of this section, we will expand on the five points above, with an aim to providing some guidelines of how one might develop this sort of fluency in a repertoire of representations. We conclude the section by presenting as an example a particular repertoire, that used in the TCIMS project, expanded with one more representation that we added to the repertoire during a trial application of the Canvas planning process.

5.1 Creation and maintenance of a representation

The first fluency that an investigator must have with a representation is the basic ability to create workproducts in that representation. This is similar to the language skill of being able to write sentences in a language. The most important thing to notice here is the relationship between the notation and the semantics of a representation. The notation refers to how something is written down; trees, tables, and lists are common examples of notations. The semantics refers to how the written workproduct is to be interpreted. There can be many different semantics for very similar notations; for example, trees are common used to reflect corporate organization (the so-called "organization chart"), class inclusion (in biological taxonomies), and heredity (genealogies).

Various tools, elicitation techniques and aids can be developed to support the creation and maintenance of workproducts in particular representations. In one Canvas follow-up project, we produced a special "taxonomy graph paper" to simplify the creation of taxonomic representations. It is also possible to have computer support for the creation of certain representations. Diagramming tools and outliners offer numerous possibilities for such support. It is important to keep in mind, however, that the semantics of the representation might not be supported by the tool. Over-reliance on the tool can result in a false sense of security that if something is represented "in the tool", everyone will agree what it means.

Since it requires something of an investment in training and infrastructure support to make widespread use of a given representation within a KA effort, it is wise to consolidate this supporting material in the KA Team Guidelines for the KA enterprise.

---

[1] In preceding sections of this guidebook we have avoided the term "knowledge engineer", using "investigator" to refer to a person who performs KA tasks. Since we now deal with a coherent set of skills necessary to perform KA in diverse settings, it is appropriate to speak of a person who has acquired these skills as a knowledge engineer (albeit with distinctions from the conventional usage of the term in the expert systems field, as explained in the lexicon entry in Appendix C).

5.2 When to use a particular representation

This is the question that most beginning knowledge engineers find the most frustrating. They have learned a small repertoire of representations; but when one is faced with a confusing mass of informants, artifacts and information to be gathered, deciding which one to use can be daunting. There are a few things that can make this decision easier.

First, consider the audience for the workproduct. As discussed in Section 3.0, there can be many audiences for a particular workproduct. Sorting out which audience will see a workproduct can reduce the uncertainty in choosing a representation. For a particular audience, what sorts of representations are they accustomed to using? We have found a good deal of prejudice (also in our own practice!) around this issue, that certain classes of practitioners are unaccustomed to certain types of representations (e.g., "only software engineers can understand entity-relationship diagrams"). Don't trust your prejudices. Do a bit of investigation on the intended audience, and see what they are accustomed to working with. Select a representation that they will be comfortable with (but beware the pitfall of the "falsely familiar" representation, described in Section 5.3).

One of the most effective ways to select a representation is based on the kind of knowledge that the representation can express. This characterization can be quite subtle, but is probably the best way for a knowledge engineer to "get to know" a representation. Below we identify some characteristics of the kinds of knowledge that a representation can express. These characteristics can have complex interactions, but they are the best way we have found to get a handle on the relative powers of different representations.

Dynamic versus Static Information

Representations vary in their suitability for representing *dynamic* and *static* information. Processes and algorithms are dynamic in nature, and require very specific notations to capture this dynamic aspect. Category structures and lexicons, on the other hand, do not have a dynamic aspect to them, and have another set of notations altogether. Dynamic and static in this sense do not refer to how quickly the knowledge changes, rather, to the nature of the thing to be explained itself. Flow charts, instruction booklets and programming languages are common examples of dynamic representations; data dictionaries, class structures, and grocery lists are common examples of static representations.

Procedural vs. Declarative knowledge

Procedural knowledge is a type of knowledge which consists of the skills or tasks a person can accomplish or perform. This is what an informant would describe as "knowing how" to do something. Procedural knowledge is often an automatic response to stimuli, can be reactive in nature, and is therefore difficult for an informant to explain or describe. *Declarative knowledge*, on the other hand, is a type of knowledge often referred to as "knowing that". It is surface level information that an informant knows he possess and that he can easily verbalize. An example of procedural knowledge and the difficulties inherent in acquiring it can be shown by considering how difficult it is to explain how to tie a shoelace. In contrast, the declarative knowledge that a shoe has two laces, that there are two types of knots that can be tied (i.e., granny and square), is easy to access.

Semantic and Episodic knowledge

Semantic and episodic knowledge correspond to two forms of long-term memory. Semantic knowledge includes words, symbols, facts, and definitions, while episodic knowledge includes

autobiographical and experiential knowledge which an informant has grouped or "chunked" together. The names of streets or sections of town are examples of semantic knowledge, while the knowledge of driving the route along these streets to get to work (e.g., the strategy for changing lanes or for operating the car) is an example of episodic knowledge.

Variability versus Commonality

How we deal with variability in a knowledge acquisition project is of fundamental importance. Variability can only be dealt with to the extent that our representations can express it. When a representation is capable of expressing variability, it means that it can express information about a set of phenomena, explicitly indicating what the exemplar phenomena have in common, and what they do not have in common. A common representation of variability is the product comparison chart, as seen in consumer information articles. A particular representation can be good at representing either commonality, variability or both; these variants, including the perhaps counter-intuitive idea that a representation might be good at representing variability but not commonality, are treated in detail in [40].

5.3 Traps and Pitfalls

Once a representation has been chosen, and some knowledge is being represented, there are still a number of problems that can arise. Each representation will have some specific problems, which are best learned by experience. However, there are some general categories of problems that can befall many different representations.

Probably the most common pitfall of a representation comes when a workproduct is shown to a practitioner in some community for the first time. The practitioner looks at the representation, and finds it confusing and incorrect. Sometimes, this can cause the practitioner to become frustrated or annoyed. Often what is happening here is that the chosen representation is very similar in notation to some representation with which the practitioner is familiar, but the semantics are somewhat different. This is a danger even (or perhaps especially!) when a representation has been chosen for its familiarity to its audience.

There are a number of ways to avoid this trap. First, if you use a representation based on its similarity to a familiar representation, pay close attention to the semantics of the representation in its native context. Sometimes the differences can be quite subtle. If you choose to use different semantics, make the differences clear to the audience. Cosmetic changes to the notation (printing diagrams rotated through 90 degrees, using different shapes for boxed information, etc.) can act as a reminder about which representation has which semantics. If the semantics continues to be confusing, perhaps it would be better to use another representation for this audience.

Another pitfall happens when a representation is pushed beyond reasonable limits. Although representations for knowledge acquisition typically have less expressive power than even most programming languages, it is still possible to use them to express a wide range of information, which might be better represented some other way. A sign of this happening is that the representation process starts needing counter-intuitive justifications (e.g., "We can represent it like this, if we just say that Lassie is a breed of dog, but with only one dog of that breed", or "this works fine, if we say that birth is the leading cause of death"). This sort of representation process usually indicates that a representation is being used in a way that will not be apparent to its readers, and hence will fail farther down the road.

5.4 Representation bias

In Section 4.8 we discussed the issue of bias in planning a thread, mainly from the standpoint of managing an investigator's bias in a knowledge acquisition session. However, human experience is not the only source of bias in a knowledge acquisition setting; the choice of a representation for knowledge acquisition also brings in assumptions and biases about the knowledge to be represented.

The notation and semantics of a representation limit what can be easily expressed in that representation. The filtering and organizing effect that this has on the knowledge acquisition process is essential to the role of a representation as a means of focusing the process. However, it also biases the representation toward interpreting the world according to the terms of the representation.

> Example. Consider an "organization chart" of a hierarchical organization. It is made up of nodes with names of individuals connected by links in the form of a tree. The semantics of the diagram specify that people further down the tree report to people above them on the tree. This representation is biased toward hierarchical organizations. It is quite easy to draw and read such a diagram for such organizations. However, if it is used in a non-hierarchical organization (say, a congregation that makes decisions by consensus, or a self-managing team), then the biases of the representation show up as difficulty in applying it.

Representational bias is a two-edged sword; on the one hand, the bias of a representation allows a particular workproduct to focus attention on desired aspects of the elicited knowledge; on the other hand, representational bias can cause an effect of "mind set", blinding the investigator (and the informant) to aspects of the knowledge that are not easily expressed with the representation. The investigator can thereby miss information not easily expressed in the representation, while over-emphasizing or even projecting information that is easily expressed.

Representational bias has an important advantage over individual or personal bias — it is objective and predictable, and hence can be compensated for. While it is difficult for a person to answer the question, "What are my biases, and in what way do they limit me?", it is relatively easy to determine the biases of a representation, and to account for them during the planning process.

These aspects of representational bias make it one of the strongest motivations for assembling a repertoire of representations. The problems of mind set can be reduced or avoided by using more than one representation to avoid getting locked into a single perspective, using the relative predictability of representational bias to coordinate the effort.

5.5 Relations between Representations

The difference between a grab-bag of tools and a repertoire is the relationship among the various tools. If one tool is not right for the job perhaps another related tool is; a job that is unwieldy using only one tool might well yield to a judicious combination of tools working in concert. We have identified a number of relationships between representations that suggest this sort of synergy.

Sloppy vs. Formal Representations

A full repertoire of representations should include a number of *sloppy* representations as well as formal ones. By "sloppy" representations, we refer to representations in which relatively few restrictions are placed on what constitutes a valid workproduct in the notation. A more formal representation is one in which the notation is restricted. These are comparative rather than absolute terms; one representation can be sloppier or more formal than another.

Example. A representation for diagrams, in which boxes are connected to each other by arrows, with some unrestricted annotation on the arrows, is an example of a sloppy representation. If the diagrams are restricted to having no cycles, and the annotations are chosen from a small set of meaningful keywords, then the representation is more formal.

The word "sloppy" is not intended to imply that these models are somehow less correct or less useful that formal models. In fact, sloppy models have a very particular use, often in the interview setting itself. During an exploratory discussion, the freedom allowed by a sloppy model can be exploited to capture rapid flow of ideas without a great deal of filtering. A formal model, on the other hand, is more likely to have specific biases, allowing the interview to have a clear focus. Sloppy models are often used to represent partial, incomplete, or uncertain knowledge, to allow for a mix of notations or semantics, and even to gather data about a given community's "native" interpretation of a given notation style.

As shown in the example above, a particular formal representation might be a restriction of a sloppy one. A common way of combining the use of two such representations is to begin general elicitation with the sloppy representation, use one or more formal representations to identify particular aspects of the information, and then follow up with a more focused, systematic investigation based on the more formal representation.

Summarizing Representations

One workproduct might represent a summary of information found in several other workproducts. Especially in a large knowledge acquisition enterprise, the representation repertoire should contain some representations that are particularly useful for summarizing others. The repertoire could include a number of different ways to summarize any particular representation, reflecting the interests of different stakeholder groups.

Complementary representations

The knowledge types outlined in Section 5.2 are described as contrasts; procedural/declarative, dynamic/static, etc. Pairs of representations that treat knowledge of contrasting types can often be used to complement each other. In the TCIMS/SEP repertoire described below, we will see a number of these complementary pairs.

5.6 A Sample Repertoire — TCIMS/SEP

As an example of the properties of a repertoire of representations, we will describe the repertoire used by the TCIMS project. Recall that the TCIMS effort was based on the SEP methodology. The repertoire of representations used in SEP starts with the *scenario*, used to describe direct experience in a work setting. Then there are several forms of *task representations*, which summarize and interpret aspects of the scenarios. As a complement to the task representations (which capture mostly procedural knowledge), SEP uses a number of concept representations for declarative knowledge. Many of the TCIMS workproducts were collected in an on-line repository called SEPWeb, from which some of our examples in this chapter and in Section 6.0 are derived. In our trial applications of Canvas, we also used taxonomies to represent aspects of variability. We have included a description of taxonomies here, although they are not a part of the SEP repertoire.

Each representation type is described and characterized in terms of the attributes of representations given above. These characterizations can be used by the reader as examples to guide in characterizing and selecting representations for inclusion in his own repertoire.

Title: Parachutist's Chute did not open

Environment: Training Jump: No enemy present

Medics and buddy approached casualty after seeing him fall

Performed initial Assessment. Findings included:
    nasal fracture (congested airway)
    patient unconscious
    Bilateral fracture — legs distorted
    Pelvis probably broken

Requested Ambulance by radio
Intubation done to improve airway
Provided patient with oxygen
Prepared patient for evacuation
Performed secondary Assessment
Straightened and splinted legs
Started IV
    patient regained consciousness
Loaded patient onto Ambulance (20 minute transport)
    patient arrived at field hospital
Field hospital performed surgery

Exhibit 11. Example Scenario from SEPWeb

5.6.1 Scenario representations

Scenarios are used to represent single threads of execution. Scenarios contain environmental context information, initial state information, actors, items in the environment, time referenced transitions, and a final state.

Scenarios can be used to describe specific situations that occurred and the events that took place in these situations. More generic scenarios can be used to describe the tasks carried out by a particular actor. There are two types of scenarios: "As Is" and "To Be." The As Is scenario documents current practice and the To Be scenario can be thought of as a proposal for the future.

Scenarios are generally presented as text files in combination with maps or other graphic information. In some cases timelines are also included. Scenarios are decomposed into events (or tasks). Each event is listed, described, and possibly further decomposed. At the atomic level, an event is a responsibility assigned to a single entity. This entity can be a person, computer, or device. For example, "Determine Pulse" is a event that could be done by a medic or a medical device. Exhibit 11 shows an example scenario taken from the TCIMS project repository SEPWeb. Notice how it describes the particular roles that the performers (buddy and medic) and devices (radios and ambulance) play in the event.

From the viewpoint of the target community, the primary uses of scenarios are to feed task decomposition (see section 5.6.2) and for system test and evaluation. From the viewpoint of the focus community, scenarios are used for validation, to help ensure the development is focused in the right direction.

Tell the domain expert: We will go through the following steps • select a scenario • describe the scenario to me as an overview • describe the actions that a situation assessment "person" would take at each point in the scenario • review the tape of the above to explicate the reasoning and justifications which you used at each point. From this I will attempt to pull out the important criteria and types of reasoning employed.

Exhibit 12. Steps for Scenario Elicitation

The development of sample scenarios that describe responses to specific situations by an informant allows us to look into a specific set of situations to get a clear idea of the tasks and responsibilities involved within the informant's area of expertise. By decomposing scenarios, we are able to clearly identify the key participants, identify the sequence of events and identify the key issues and communication requirements generated during the scenario. The informant and investigator should work together to determine the major occurrences of each scenario.

Providing a skeletal structure for each scenario developed ensures that the key elements are identified. Exhibit 12 shows an example of structured steps that can be used for scenario elicitation. Working from this set of steps, the investigator and informant further define the scenario and its attributes. Scenario elicitation of this sort is an example of collaboration between the informant and the investigator. As an informant gains familiarity with the scenario form, he becomes more adept at providing scenarios. In this way, the scenario elicitation technique is training the informant.

Scenarios are a way of capturing procedural knowledge in a specific situation. Like processes, scenarios describe dynamic aspects of the world. Variability in the scenario is not ordinarily an issue, as a scenario represents a single thread of execution. Variability in the domain of the focus community is addressed by composing a set of scenarios that encompass the important aspects of the domain with sufficient variations to be adequately descriptive.

5.6.2 Task representations

Task representations allow decomposition of a task into its subtasks. Task decomposition can be used to represent a task performed by the focus community that occurs repeatedly over a wide range of scenarios. For example, the task of "Casualty Assessment" occurs repeatedly in all battlefield care scenarios, although when it is performed (and by whom) differs.

Task representations are particularly useful in the selection of areas to be considered for automation (using the As Is decomposition) and in determining how automation will be accomplished during system development (using the To Be decomposition); therefore they are often the representations of choice in knowledge acquisition projects aimed at technology development. Task

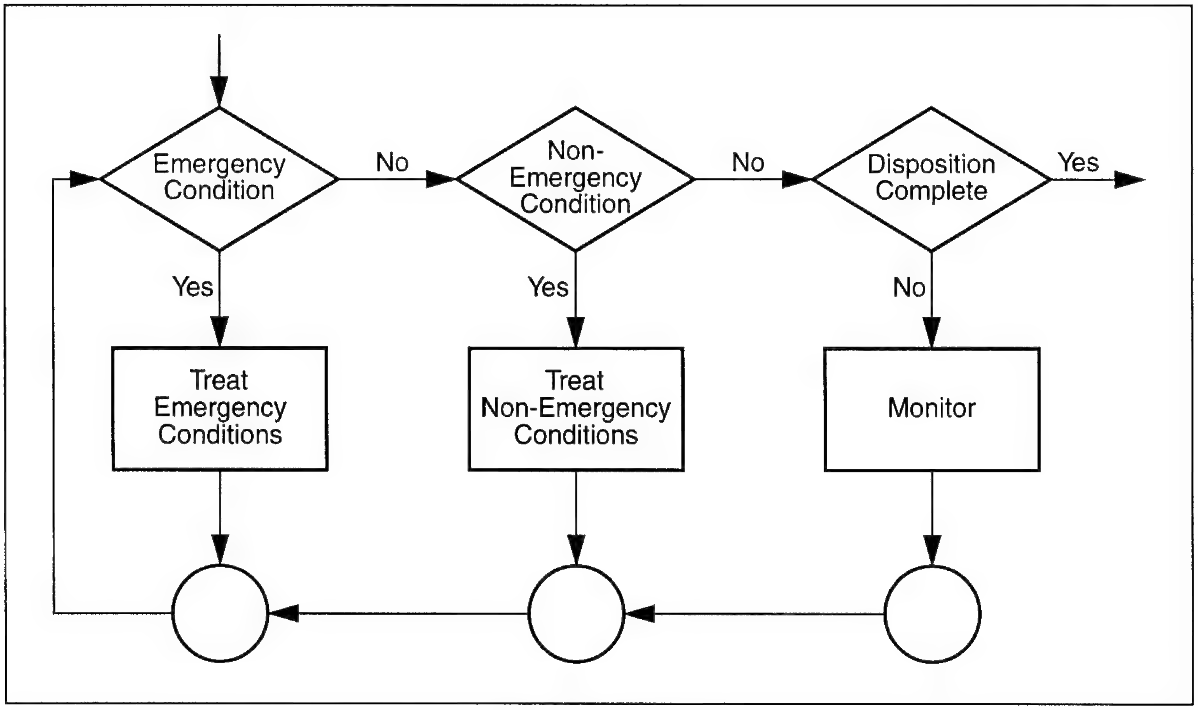


Exhibit 13. Flowchart

analysis is used to identify major tasks or job functions and their relation to the overall job being performed. This type of analysis requires the investigator to be able to break tasks and subtasks down into manageable chunks that will be represented.

The task representations in the SEP repertoire vary widely, and can be selected on the basis of the character of the tasks being described. The SEP repertoire includes simple textual descriptions (not shown) to capture information about subtasks, timing, constraints, decisions, and resources. The repertoire also includes some more formal graphical representations, namely flowcharts, modified petri nets, event trace diagrams, and state transition diagrams. Flowcharts are best suited for representing decision-oriented task sequences (see Exhibit 13). While flowcharts are particularly biased toward simple linear sequences, modified petri nets and event trace diagrams can be used when interaction and coordination is a dominant element. The modified petri net (shown in Exhibit 14) is a petri net with the additional features that tokens may be treated as objects that retain their identity through transitions. The modified petri net is well suited to provide a view of overall interaction and coordination, but lacks temporal information. Event trace diagrams (shown in Exhibit 15) can easily represent time, but the complexity of the diagram grows quickly as its scope is increased. As appropriate, state transition diagrams can be used to record transitions associated with performing tasks (not shown).

Task representations are versatile in the types of knowledge they can represent. While they primarily represent procedural rather than declarative knowledge, they can express both semantic and episodic knowledge. Like scenarios, tasks represent dynamic processes. The degree of representation of variability in task representations varies greatly. The event trace diagram is essentially synonymous with scenarios in that it shows a fixed sequence of events, so there is essentially no variability present except any variability in times associated with each task component. A flow chart represents the limited variability supported by decisions within the process being modeled. The inherent non-determinism of the petri net implies that any single diagram represents a broad range of behavior subject to the constraints imposed by the coordination activities, such as "Vehicle Management" in Exhibit 14.

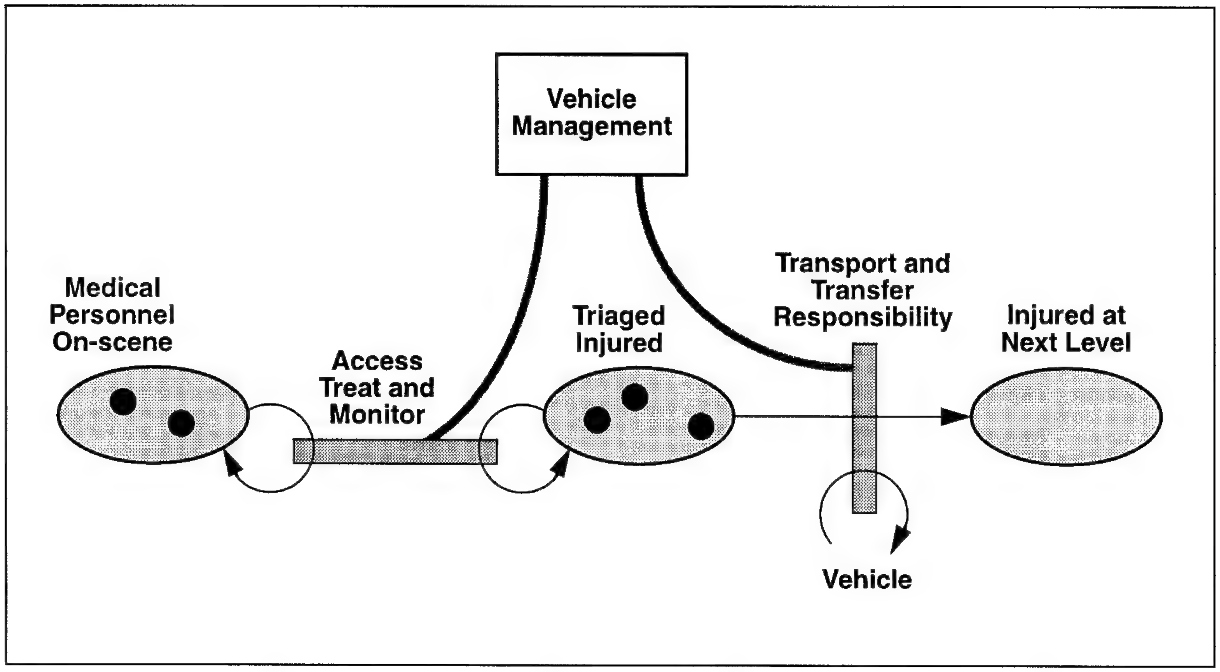


Exhibit 14. Modified Petri Net

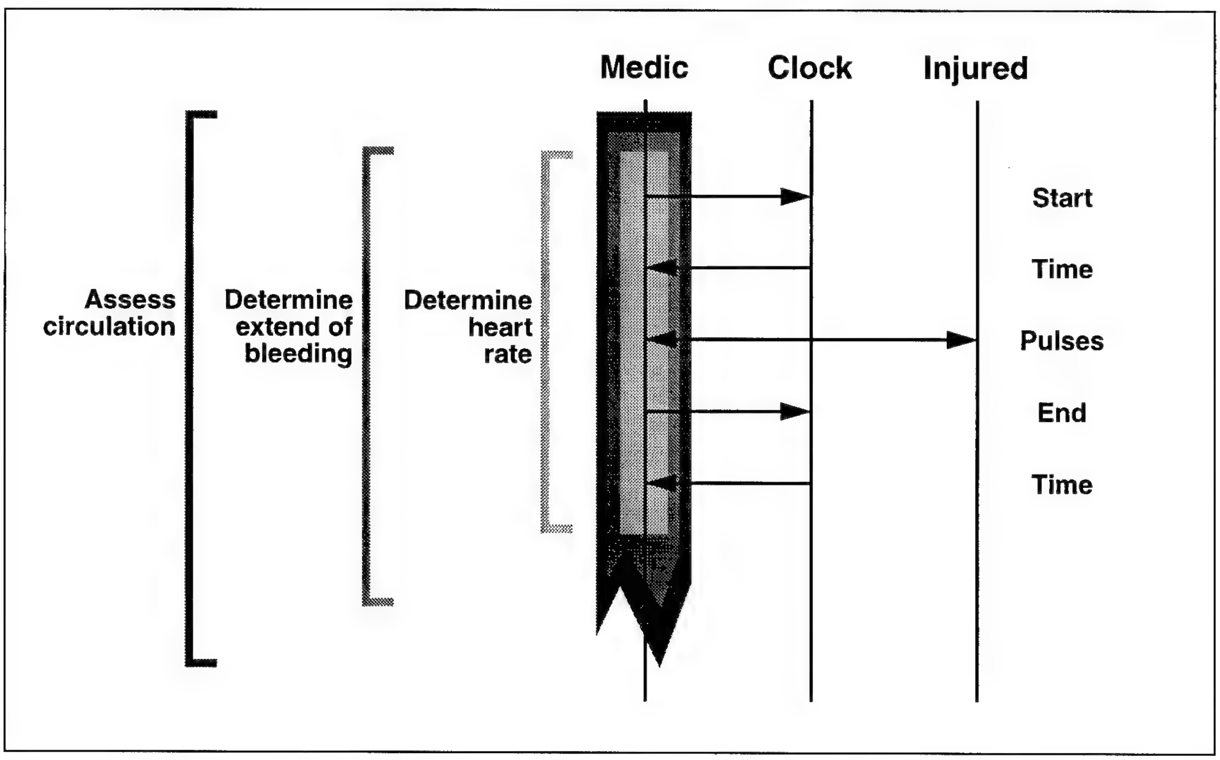


Exhibit 15. Event Trace Diagram

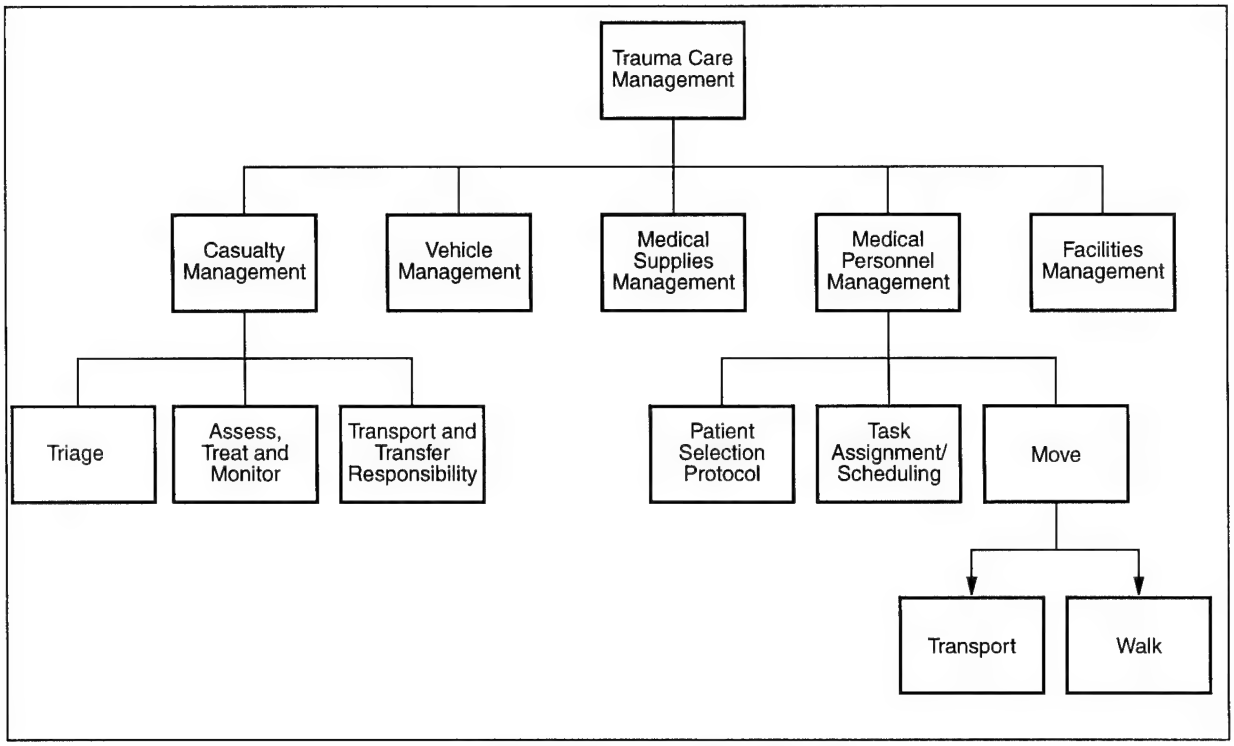


Exhibit 16. Conceptual Hierarchy

5.6.3 Concept representations

Concept representations are used to differentiate between concepts identified by an informant. If the audience, at any level, is unfamiliar with concepts being represented in other representations, conceptual analysis and representation is vital to a project's success. If the audience is familiar with the domain and its terminology, the need to identify and represent primary concepts is reduced. Providing concept representations is also critical when the audience is diverse, such as a consortium rather than a single company. Identifying and modeling domain concepts allows all participants to understand and agree on common domain terminology.

Graphical representations that can be used to represent concepts include conceptual hierarchies, concept maps, and object diagrams. Exhibit 16 provides an example of a conceptual hierarchy. Conceptual hierarchies show aggregation of concepts. This type of view is effective when the domain is large and concepts must be consistent for presentation to a large audience.

Another means to represent concepts is the concept map, shown in Exhibit 17. This representation is created by having an informant describe the various attributes within a familiar domain. To create the concept map shown in Exhibit 17, the informant, an emergency medical technician, was asked to describe pre-hospital emergency medical providers. This map was then used to further define attributes and tasks within the overall medical scenario.

A somewhat more formal representation for representing concepts is the object diagram with relations, shown in Exhibit 18. This representation is most appropriate when the audience includes a software development community of practice. Object maps can be used to elaborate a concept map when the concepts involved form a taxonomic classification.

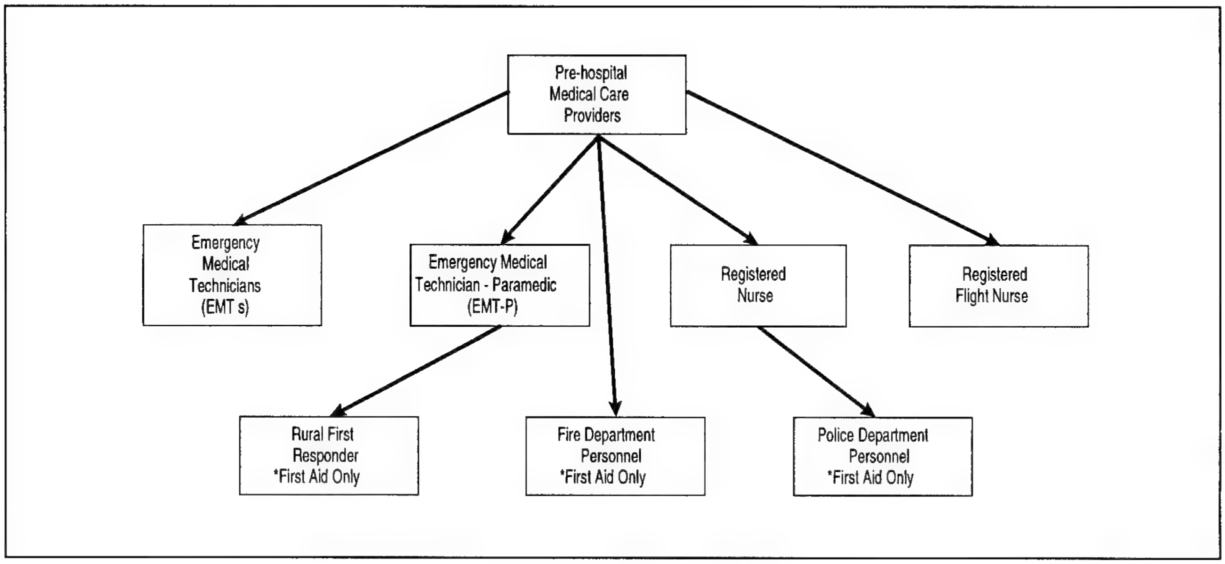


Exhibit 17. Concept Map

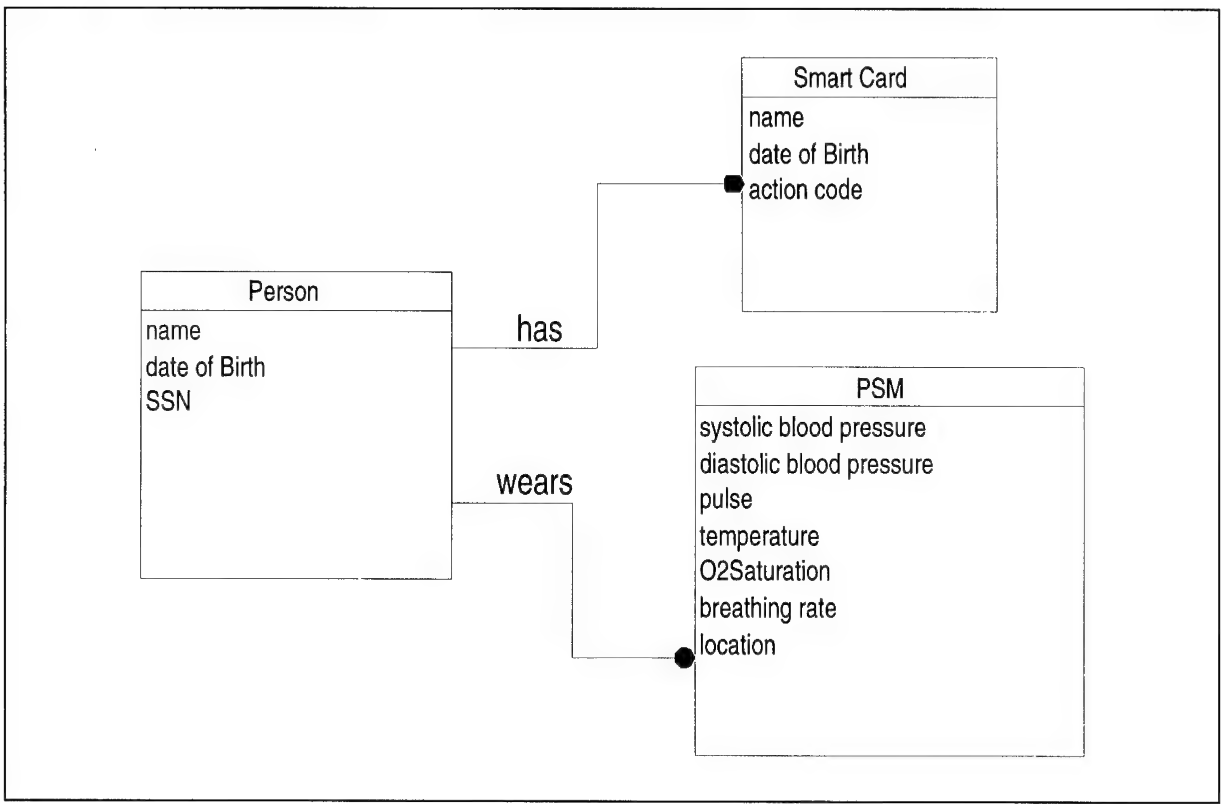


Exhibit 18. Object Diagram with Relations

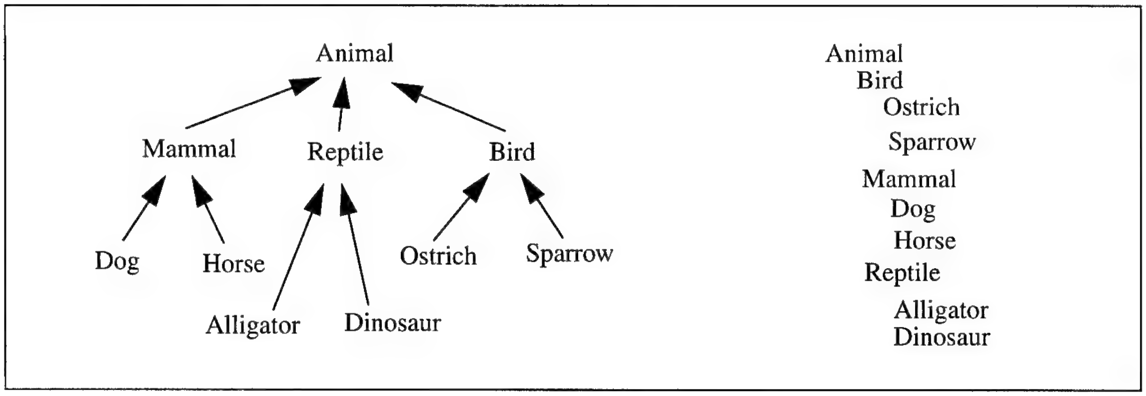


Exhibit 19. Two Ways of Notating the Same Taxonomy, as a Tree and as an Outline

Concept representations differ from task diagrams in that they concentrate on declarative, rather than procedural knowledge. They share with task diagrams the ability to represent both semantic and episodic knowledge. Typically, concept diagrams represent static entities, like categories or objects and their properties.

5.6.4 Taxonomy representations

A taxonomy is a hierarchical representation; its notation is any of a number of hierarchical notations, as shown in Exhibit 19. The semantics of a taxonomy state that the relationship between nodes is *specialization* (i.e., more specific concepts appear below more general concepts).

Taxonomies are hierarchical, as are parts breakdowns, organization charts, status relations (pecking order). Almost every field of study uses hierarchies. The most common pitfall in the use of taxonomies is to assume thereby that everyone understands them the same way. However, this familiarity of hierarchies can lead to problems in taxonomy interpretation by informants, since the semantics behind a taxonomy are fairly abstract. In a given knowledge acquisition session, if the informant has a high degree of interest in one kind of hierarchical relationship and is shown a taxonomy, the informant may conflate the semantics of the taxonomy with the semantics of their hierarchy of primary concern.

For example, if a medical practitioner sees the concept "surgeon" under the concept "physician" (meaning, according to the semantics of the taxonomy, that a surgeon is a special case of physician), she might misinterpret this to think it says that a surgeon is less qualified than a physician, or that surgeons report to physicians.

The warning to take from this when planning knowledge acquisition is that it is important to understand the interpretation placed on hierarchies by the intended audience of a taxonomic representation. It is always the case that the choice of a representation will constrain the choices for a workproduct audience; in the case of taxonomies, the ubiquity of hierarchical diagrams makes this problem particularly subtle.

Taxonomies, like concept diagrams, are typically used to represent declarative knowledge. Taxonomies can represent static relationships between entities, and are not well-suited to representation of dynamic information. However, taxonomies do explicitly allow the representation of variability through different specializations of a single concept. When combined with relations among the

concepts (as is done by modeling frameworks such as RLF [53], [54]) the details of the variation can be expressed. Variability can be expressed at several levels, indicating successive refinement. In Appendix B of this guidebook, we demonstrate how a taxonomic modeling method can be used as a representation for representing the variability in a knowledge acquisition plan itself.

5.7 A Repertoire of Representations

At the beginning of this section, we stressed the importance of a repertoire of representations, as opposed to reliance on a single representation to cover all cases. We have also shown a particular repertoire used by a particular mature knowledge acquisition methodology (SEP). This repertoire has been developed in a particular context, and has its own assumptions about the nature of the knowledge acquisition task to which it is particularly suited. Some of these assumptions have been made explicit, but it still requires considerable work for a new SEP practitioner to gain fluency in the repertoire. In addition, the repertoire presented here contains at least seven different representations; gaining mastery in all of these, as well as the fluency required to decide which one to use, requires extensive practice and experience.

It is for this reason that we encourage new knowledge engineers to begin collecting their own repertoire. In this way, they can develop their own understanding of the relations between the representations in the repertoire, and the assumptions about where they can be used. They should, of course, feel free to take any of the representations from this section and continue to add to their own collection. In addition, there are many useful representations for the initial capture of knowledge that are not complex in terms of their basic structure. The important thing is to understand the relationships between the collected representations, to know what each should be used for, and when to use it.

For experienced knowledge engineers, we suspect that they have already collected a repertoire, even if they have not thought of it in that way. We hope that the criteria and examples in this section can help them to organize and broaden their repertoire, and gain some insight into more effective ways to use it.

6.0 Dossier Planning and Management

Management of the ***KA dossier***, the repository where all materials produced or used in the knowledge acquisition effort are stored, is central to managing the overall knowledge acquisition effort. The dossier contents can include the original workproducts from the focus setting (manuals, textbooks, legacy code, screen dumps, demo materials, program documentation etc.), as well as derivative workproducts produced by the knowledge acquisition team. These could be videotapes of interviews, reports based on notes taken by the investigator during an interview, high-level diagrams such as those outlined in Section 5.0, lab notebooks (raw data) from experiments etc. Although the dossier could very well include a large amount of off-line material, its index could still be supported by a on-line indexing mechanism.

The dossier therefore contains a large number of highly disparate workproducts. In order for these workproducts to achieve the goal of knowledge acquisition, namely, to transfer knowledge from one community of practice to another, information about the context of the artifacts and workproducts must also be preserved. The elements of knowledge acquisition (Section 3.0) and the planning process itself (Section 4.0) have been designed to draw attention to this contextual information. The major role of the dossier is to record this information, for use both during the knowledge acquisition process and after the knowledge acquisition effort has finished.

More specifically, the dossier plays the following roles:

- *Serves as a delivery vehicle to the audience.* The dossier stores the workproducts that are to be read by the audience of the KA effort. Every workproduct has a specific audience for whom it is intended; the value of the KA effort hinges on how well the audience can find and understand the workproducts that are intended for it.
- *Directly records the artifact threads.* Of the various types of threads that make up the of the knowledge acquisition "canvas", informant and investigator threads are abstractions that have to do with human beings' state of mind and hence cannot be stored directly. However, the artifact thread (as described in Section 3.2.3) consists of a series of annotations to a workproduct. Each of these annotations will be stored in the dossier; a well-structured dossier will show the heritage along this thread, allowing future plans to reference it, to determine what has already been done under what circumstances, etc.
- *Provides source materials for improving investigators' knowledge.* A KA effort with a large number of investigators must carefully manage the gaining of knowledge by these investigators. As the effort progresses, more and more knowledge specific to the KA effort will be available, and the informants who are most deeply involved will come to expect that this information can be used as a base for further discussions. The well-structured dossier will also serve as a source of training material, where new investigators can catch up to become productive, or veteran investigators can keep abreast of one another.
- *Reflects planning criteria.* Certain criteria for selecting elements of the planning process reflect contextual information about the workproducts. These elements can be used to structure the dossier, so that this contextual information will be reflected there. In particular, the dossier needs to be indexed around the following:
  - — The audience for which the workproduct is intended,
  - — The role that the knowledge source (e.g., informant or artifact) plays in the focus community, and
  - — How the knowledge is represented.

In Section 6.1, we provide detailed guidance for creating an index for a dossier. In Section 6.2, we show some sample scenarios for how to use a dossier index that was created in that way; finally, in Section 6.3, we discuss possibilities for automated support for searching the index to support those scenarios.

6.1 Structuring the Dossier

The index to the dossier will have the form that we will call a *multiple hierarchy*. The easiest way to describe a multiple hierarchy is by comparison to a cross-reference index to a book of poetry. Suppose that the book has an author index, in which authors are clustered into "classical" and "modern", with modern authors further clustered into "American" and "English". Suppose further that the book has another index, alphabetical by title of poem, and still a third index, alphabetical by first line. Each of the indices also gives page numbers for the poems.

The author index is a simple hierarchy in the sense intended here for the index of the dossier. In the case of the poetry collection, the other two indices are not hierarchical. A multiple hierarchy is a compound index of this sort, where all the indices are hierarchical. Each hierarchy organizes the material based on some different aspect of the workproducts (just as the three indices of the collection of poetry organize the poems based on different aspects of the poems).

The index to the dossier will be a multiple hierarchy, with one hierarchy for each planning element that reflects the context of the workproduct in some way. In this section, we will present what we call *starter sets* to show how to construct these hierarchies. The hardest part of constructing a hierarchical index is often getting started; the idea behind a starter set is that, based on some general principles governing the index, we can prepare the top few levels of the hierarchy, which are likely to be quite similar for any dossier. As part of the knowledge acquisition planning process, these starter sets are then modified and extended to cover the details of a particular knowledge acquisition effort. For each starter set, we provide a set of questions to guide the extension process. Some are simple yes/no questions that determine whether a particular category should be included. Others are more open-ended questions designed to elicit further categories.

In this section, we will provide starter sets for hierarchies based on:

- the intended audience;
- the knowledge source; and
- the selected knowledge representation.

We have chosen this set for three reasons:

1) A great deal of information is known about these categories at planning time, allowing them to be used for the initial structure of the dossier.

2) These elements are particularly relevant to the capture of the contextual information that is critical for the successful use of the dossier.

3) The hierarchies corresponding to these elements are sufficiently similar from one dossier to the next, that it is practical to provide a starter set for them.

In order to use these hierarchies to index workproducts, information about the audience, the knowledge source, and the representation must be tracked for workproducts produced as part of the knowledge acquisition effort. The dossier can also be used to index the artifacts used as source materials (manuals, guidelines, etc.); however, this information will not necessarily be available

for such artifacts. In this case, the index can offer some guidance about what information should be gathered about a particular artifact.

As part of the TCIMS knowledge acquisition effort, a large dossier of derivative knowledge acquisition workproducts was created under the name SEPWeb. After each starter set, we will illustrate how it can be refined to form a working index of the materials in SEPWeb. Since the SEPWeb is still in development, we do not attempt to make a report of the state of the search capabilities of SEPWeb at any one point. The diagrams in this section are indices based on the publicly available data in SEPWeb, and were constructed as illustrations of the principles outlined above. The examples here have been simplified from the SEPWeb and TCIMS material, both for illustrative purposes, and to protect TCIMS consortium confidential information.

6.1.1 Audience

Exhibit 20 shows a starter set for an index of audience types. The top level is taken from the overall model of the knowledge acquisition session that has been used throughout Canvas.

In order to customize this model for your project, the following set of questions might be useful:

- Is the goal of the KA effort to provide new technology for the focus community? If so, then there are probably one or more "Implementer" audiences in the project, as well as one or more "User" communities in the focus audience.
- Is the goal of the KA effort to change work practice in the focus community? If so, then some group of "Policy Makers" will be an audience for the project.
- Is the goal of the KA effort to return information to the focus community? If so, then a "Student" community will be an audience for the project.

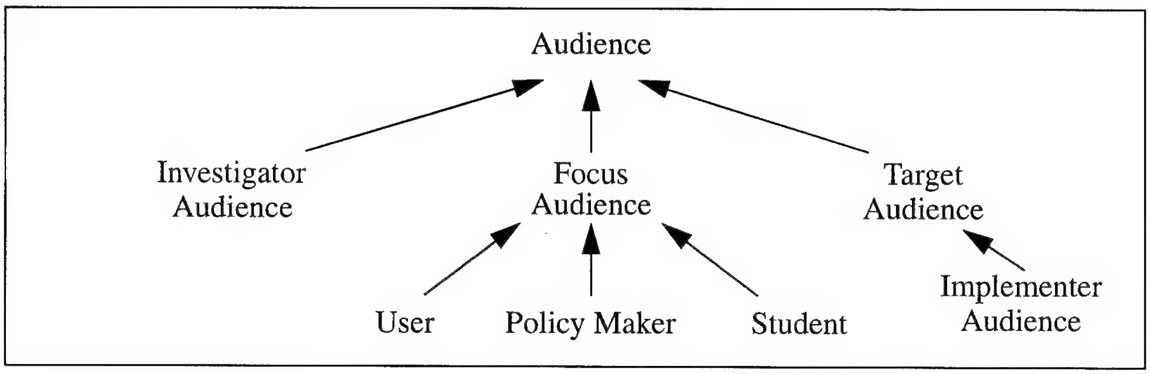


Exhibit 20. Starter Set based on Intended Audiences

The starter set shown here is intended to be suggestive of where boundaries between different audience groups are often found. For example, the term "Users" in Exhibit 20 assumes (as indicated in the questions listed above) that the goals of the project include providing new technology for the focus community, and refers to practitioners in the focus community who perform activities that will make use of this new technology. These could include virtually any kind of practitioners (e.g., secretaries, accountants, researchers, developers, etc.). There will often be several user groups, possibly with further specialization relations.

The term “Policy Makers” refers to people who make decisions about the work setting: directors, technical advisors, and project managers are typical examples. The set of indices based on distinctions made in the focus audience is likely to develop as the project proceeds, and new categories of practitioners are identified, or new distinctions of work setting are discovered among known groups.

> Example. Exhibit 21 shows the index of audience communities in SEPWeb. In this case, the greatest variety of communities was found within the focus community. Following the set of questions above, we find that one goal of the project is to change the work practice in the medical and military communities (by introducing new technologies there); therefore, these two communities appear in the index. The questions also indicate that policy makers should be included. In this case, the policy makers are the people who perform administration of the clinics. Within this group, we have another level of distinction between the clinic directors and the receptionists who keep track of the everyday business. Following the third question, since we are interested in returning information to the medical part of the focus community, medical students are also included.

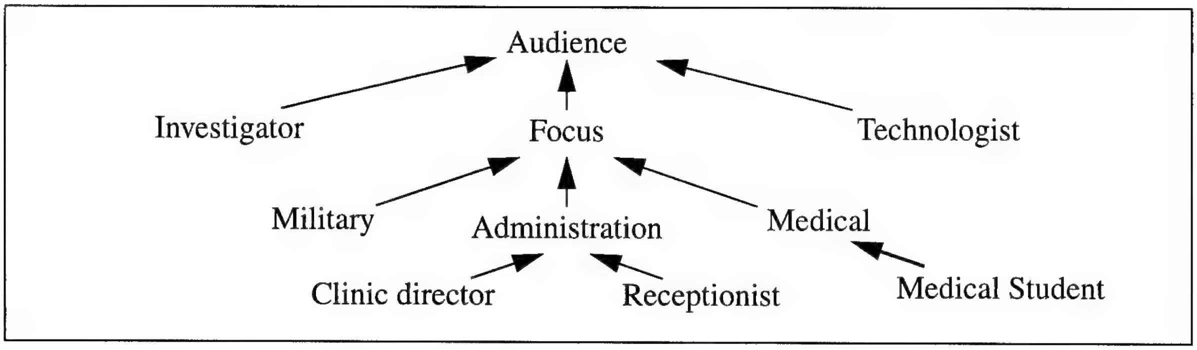


Exhibit 21. SEPWeb Index based on Audiences

6.1.2 Knowledge Source

Exhibit 22 shows a starter set for an index based on knowledge sources. The two main categories of knowledge source, informant and artifact, form the basis of the index. To refine these categories further, the guiding principle is to consider types of knowledge source that are likely to shed a different viewpoint on some topic, and hence could be used to help someone find the resulting knowledge acquisition workproduct, when he shares that viewpoint.

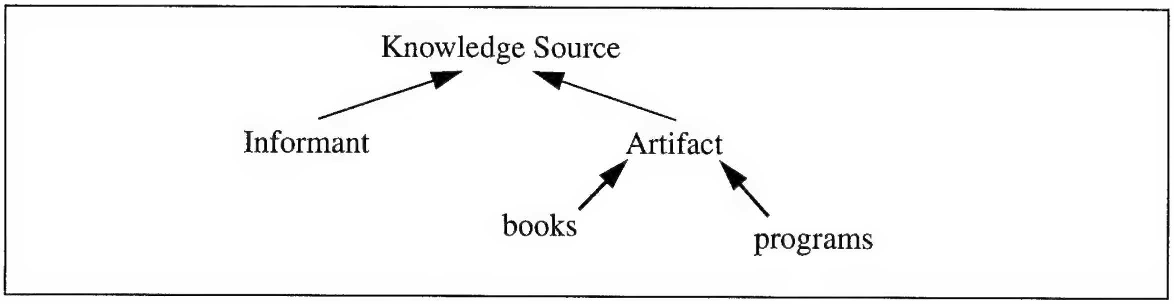


Exhibit 22. Starter Set based on Knowledge Sources

The following set of questions can help to identify informant groups with differing viewpoints. Much of this information can also be found in the community of practice view described in Section 4.3.1:

- Given one community of practice, with what other communities do its practitioners commonly interact? Builders and electricians are an example of communities of practice that relate in their professional activity.
- What are some well-known conflicts between communities? Copy editors and authors are an example of this sort of division.
- Is there professional stratification within the community? Doctors and nurses are an example of communities separated by professional class.

This starter set is quite open-ended, and can be extended with an elaborate structure, depending on the details of the focus community. The first few steps of such an extension are shown in the example below.

Example. Exhibit 23 shows the index of knowledge sources for the workproducts in the SEPWeb dossier. The starter set shown in Exhibit 22 suggests that we should consider the artifacts that were studied as well as the informants. However, since TCIMS was primarily an interview-based knowledge acquisition effort, there are few formal reports based upon artifact studies.

On the other hand, several distinctions among informants were made, only a few of which are shown here. The first distinction, between "medical" and "transport," exemplifies distinct communities that interact professionally in the TCIMS domain; in particular, medical personnel often have to be transported to accident sites, and to be transported with patients safely back to a medical facility. Transportation personnel have to be able to interact with the medical personnel in order to facilitate this transfer.

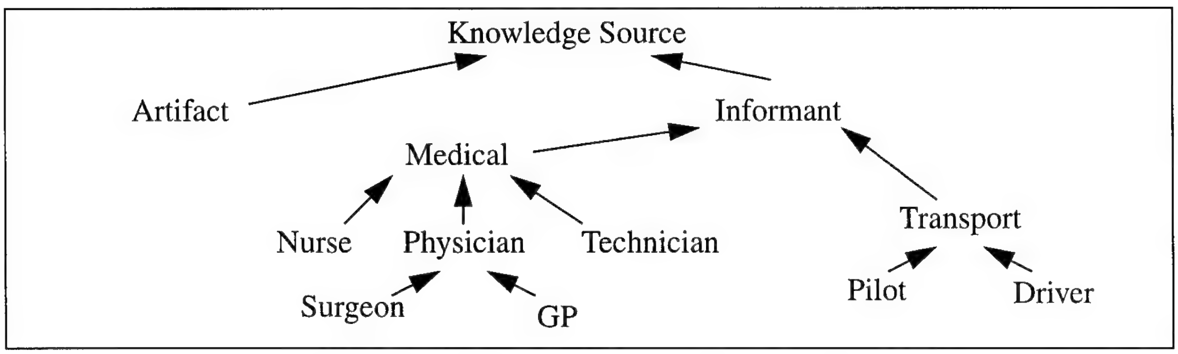


Exhibit 23. SEPWeb Index based on Knowledge Sources

Within the medical community, the professional stratification of physicians and nurses is recorded in the separation of nodes for those two groups. From these two groups, following the questions above, we find the community of lab technicians, with whom both of these commonly interact. Finally, surgeons and general practitioners are separated within physicians, because of the difference in viewpoint that these two groups often have.

6.1.3 Knowledge Representation

Exhibit 24 shows a starter set for an index hierarchy of the major types of knowledge representations, as described in Section 5.6. These are based on the TCIMS/SEP repertoire; in order to customize this to another KA effort, the index should reflect the representations in that repertoire. If, in addition to knowledge acquisition workproducts, artifacts are to be indexed by this hierarchy, then representations used in the artifacts should also be included here.

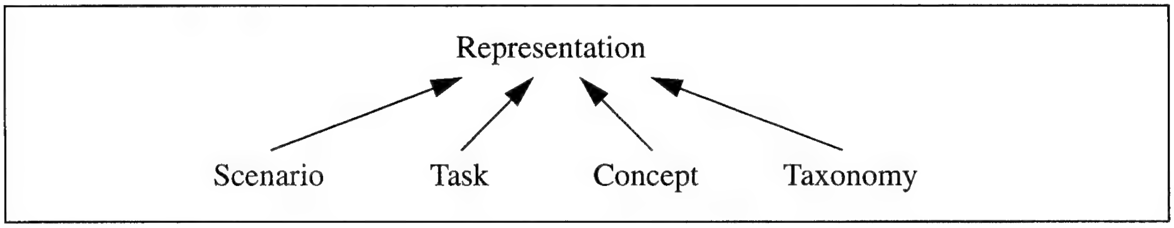


Exhibit 24. Starter Set based on Knowledge Representations

The knowledge representation choices identified in this hierarchy should be based on the decisions made during KA planning, when the representations were matched to the audiences (Section 4.6). The attributes mentioned in Section 5.1 can be used to select representations; they are summarized in the set of questions below:

- Will procedural or dynamic knowledge be collected during the project? Procedural and dynamic knowledge is usually best represented using scenario or task representations.
- Will information be available about specific situations and events? If so, it can be represented using a scenario representation.
- Is the project interested in hierarchical classification of information? If so, the hierarchical classification can be represented using a taxonomy representation.
- Does the focus community already use hierarchical notations? If so, then the dossier should be organized in such a way that the different hierarchical notations will be clearly distinguished.
- Will there be static knowledge of the attributes and relations among entities? If so, then the static knowledge can be represented using concept representations.

This set of questions should be extended, along with the starter set, for the representation repertoire being used in the knowledge acquisition effort. The following example shows how this index has been expanded in the case of SEPWeb.

> Example. Exhibit 25 shows the categories of representation notations used in the SEPWeb. Each dossier entry is labelled with the representation used to express it. Following the set of questions given above, since TCIMS concentrates on acquiring process information, both in the case study form (scenarios) and general form (tasks), these two representations appear in the SEPWeb index. In TCIMS terminology, "object representation" is the name given to a taxonomic model. TCIMS is a large, comprehensive effort, so it includes all the types in the starter set, along with a further type, the flow representation, for recording coordination information among the many actors in a military medical situation.

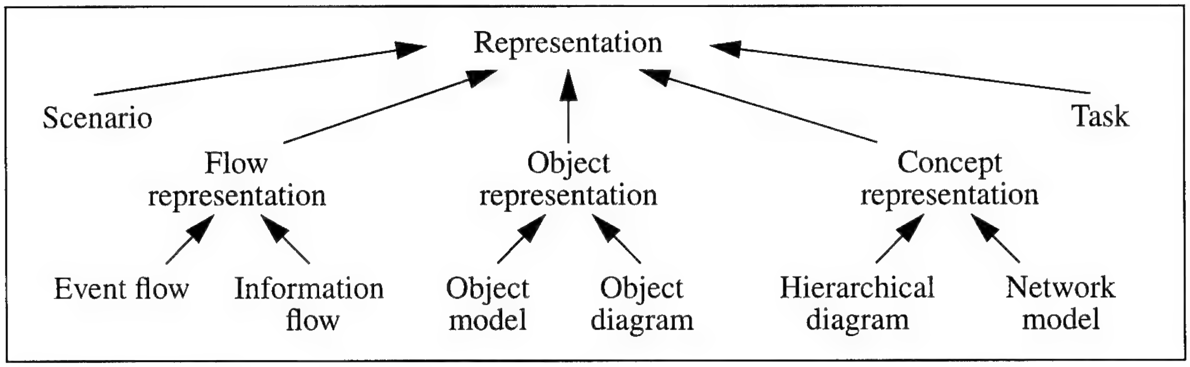


Exhibit 25. SEPWeb Index based on Knowledge Representations

6.1.4 Topic

Although one of the most powerful methods for indexing a dossier is by topic, we have neglected it so far in our list of starter sets. The reason for this is that, unlike the audience, knowledge source, and representation indices, topic indices vary greatly from one dossier to the next; thus there is little or no commonality from which we could build a starter set.

Recall that in Canvas, the word ***topic*** refers particularly to something known to the focus community that is the focus of attention in some KA session. No matter who is consulting the dossier, it is likely that they will have some idea of what topic they are interested in learning about. A particularly useful option for managing a dossier based on topic is to use keyword searches, since the keyword lists can easily be updated along with the knowledge acquisition workproducts themselves. In a large dossier like SEPWeb, this can lead to problems when the set of keywords becomes too large. SEPWeb solves this problem by structuring the set of keywords into three sets, one each for the scenario representation, the task representations, and the other (flow, object, concept) representations. Further structuring of the keywords is possible, and is the subject of further research for the SEPWeb team.

In our pilot application of the Canvas planning process, we found that a hierarchical organization of the candidate topic areas can also be useful, especially for initial baseline sessions that are intended to provide an overview or "roadmap" of the broad area of investigation. This kind of hierarchical breakdown of topic areas would, of course, provide an excellent basis for this facet of the dossier index; however, the actual content of the hierarchy will be highly project-specific.

Summary

The starter sets presented in this chapter were chosen as a useful illustration of how the parts of the multi-hierarchical index can be formed. In any particular KA effort, other hierarchies might also be useful, based on any aspect of the planning process. For example, if the effort involves many different types of sessions (individual interviews, group sessions, workplace observations, etc.), then an index based on session type could be useful. If there is possibility for confusion in the language or terminology used in the workproducts, an index of the different communities of practice according to language (similar to the audience index discussed here) could be useful. In short, any information about the context of the workproduct should be tracked, and can potentially play a part in one of the indices.

6.2 Sample Usage Scenarios

The structure of the dossier can be used in many ways to support processes involving the use of the dossier. These range from planning the knowledge acquisition process itself, to access by end users of the dossier. We will outline a number of scenarios, and show how they make use of the indices outlined above. The scenarios given here are only the tip of the iceberg; many other scenarios are possible. All the scenarios have one thing in common — they respond to some need to understand the workproducts in the dossier, in the context they were created or used.

6.2.1 Use in Managing the Ongoing KA Effort

The elements in the dossier can be used during the knowledge acquisition effort to manage all of the threads. Some example uses are as follows:

- *Perform a "walk-through" of a KA workproduct with a new informant* (i.e., not the original knowledge source for the workproduct). To acquire more detailed knowledge, an investigator can show KA workproducts to another informant, and elicit commentary. To do this, he needs to use the audience index to verify that the new expert is in the intended audience, or to handle the resulting difficulties if he is not.

  Example. We might want to ask nurses to review reports of procedures given by physicians, to see what their viewpoint can bring to the picture. We can use the audience, knowledge source, and representation indices to support this plan; we use the knowledge source index to determine if the workproduct came from an interaction with a physician (including surgeons and general practitioners), we use the representation index to restrict it to task representations, and finally, the audience index to make sure that we are using a diagram that was intended for a medical audience, so that we can reasonably expect the nurses to understand it.

- *Perform a comparison or correlation analysis.* We might want to study several KA workproducts together to find correlations or comparisons. We can use all the indices, depending on what we want as the common theme of the comparison.

  Example. A systematic comparison of nurses' and physicians' views on process could yield an interaction diagram that shows where these two communities have to negotiate some limited resource. Such a comparison would make extensive use of the knowledge source index.

- *Have a new investigator interview an informant who has been interviewed before.* When a new interviewer wants to follow up another interview, he might want to read up on the reports that were elicited from the same informant before, either to track the informant's changing opinion on a topic, or to avoid eliciting the same information from the same person twice. This usage corresponds to tracking the evolution of the project along an informant thread.

  Example. A physician might have modified his view on a particular procedure after seeing the comparison between the nurses' and physicians' viewpoints. Investigators pursuing further interviews with this physician should be aware of this history. The knowledge source index can be used to find the relevant workproducts.

- *Prepare an investigator to interview a new informant.* If an investigator wants to interview a new informant, he might want to "do his homework" so that he can maximize the benefit of the time with the informant. This could be achieved by examining other knowledge acquisition workproducts that are already available from sessions with other related knowledge sources, such as for the work setting where the informant is a practitioner. This usage corresponds to evolving the project along an investigator thread.

Example. Suppose that a large KA project gains access to a high-profile expert to use as an informant. She is the only known source of information about the effects of residual radiation on military installations; however, in order to understand her theory, one needs to know a number of details of how military installations deal with hazardous materials in general. The planner decided to have the interviewer prepare by examining all the relevant workproducts in the dossier.

6.2.2 Intended Use by Target Audience

The dossier could have a number of intended audiences, managed by the audience index described above. When the dossier is used by a member of one of these audiences, it can be used to do the following:

- *Gain an overall view of some topic.* In this case, the user would want to find all workproducts on a given topic, of any representation type, intended for any audience of which he considers himself a member.

  Example. If a clinical director wants to have the big picture about blood supplies, he can use the audience index to filter the workproducts to find those that are intended for himself as an audience, and use the topic keyword list to find the workproducts relating to blood supplies.

- *Find out what practitioners in some setting say about some topic.* The user can consult the sources index to find the workproducts that were produced by some other practitioner group.

  Example. A physician interested in knowing what lab technicians have to say about data confidentiality can filter using the knowledge sources index to find only those workproducts.

- *Locate a specific report.* All the indices can be used together to find a particular report.

  Example. A system implementer is planning the specifications of a new program to be used by the BAS surgeon. She wants to know all the interactions that the BAS surgeon has with other agents. She begins with the BAS surgeon from the keyword list; this results in several workproducts. Since she knows that she is interested in relationships between entities, she can restrict her search to concept representations. She finally finds the workproduct she needs under the network model of the BAS surgeon.

6.2.3 Future Spin-off Uses

By tracking the context of the workproducts in the dossier, and by recording whatever information about the context of the artifacts used in the knowledge acquisition effort, the dossier can provide value beyond the original stakeholders identified at the beginning of the effort. A future audience can recover whatever information about the context that is important for their purposes.

Example. A medical student working with one of the physicians interviewed in the TCIMS project decides that he would like to use the dossier as source material for a class. He can use the audience index to determine which knowledge acquisition workproducts were intended for a medical audience (since those intended for other audiences might have inaccuracies or imprecisions that are of concern only to medical practitioners). Similarly, the knowledge source index can help him to determine the reliability of the particular information.

6.3 Possibilities for Automation

The usage scenarios given above suggest ways in which the dossier could be indexed, so that knowledge acquisition workproducts can be stored and retrieved in useful ways. In order to gather the information needed to form these indices, the knowledge acquisition sessions must keep careful track of the information used in workproducts, in particular, the source of the information, the intended audience of the write-up, the representation that is used, etc. The exact details of which information will be included in the indices should be determined during planning of the enterprise, as described in Section 4.0. The KA workproducts must then be stored according to these indices, and made available to the entire investigator team. Given that this team might well be widely distributed geographically, this distribution is a perfect opportunity to use the capabilities that are provided by the technology of the Internet.

The wide variety of types of information available to index the dossier provides a wide range of opportunities for automatic support. In this section, we outline these possibilities, some of which are already under development.

6.3.1 Web Accessibility

Many of the requirements of a knowledge acquisition dossier are similar to those found on the World Wide Web (WWW), namely, that a large corpus of information, divided into separate pieces, needs to be viewed in a flexible and exploratory way from geographically diverse locations. Entries in the dossier are related to one another in many different ways. This suggests that we might use some of the technology already available on the WWW to automate the dossier.

One capability that has already been tested with SEPWeb is to provide a search engine for a set of dossier entries, which can find an entry based upon keywords. If we associate keywords with topics, this can be an effective way to find a workproduct related to a given topic. However, often a user sees a topic mentioned in a workproduct, and would like more information about that topic. SEPWeb already implements hypertext links for such references in text-based workproducts. For diagrammatic workproducts, the user could return to the search engine and search for the new topic, but far better would be a hypertext link from the original workproduct diagram to a list of workproducts about the related topic. The problem with this approach is that it involves constructing a hypertext active display for each workproduct, and linking it to its related topics. For general diagrams, this could incur considerable expense.

This expense can be reduced by providing a set of standard representational notations, along with web-aware tools for representing knowledge using these notations. One drawback of such an approach is that it limits the range of notations that can be used. A far more serious drawback is that the notations must be specified sufficiently formally that they can be automatically processed to create the hypertext diagrams. Many representational notations (especially "sloppy" representations, as described in Section) get much of their power from their informal flexibility. It is not uncommon to mix natural language with an approximation to a more formal notation (such as entity relationship diagrams or flow charts) in a single representation; requiring that the representations meet some notational standard could impair this flexibility.

6.3.2 Hypertext Organization of Indices

One way to make use of hierarchical indices such as those given in this section is to provide the hierarchical structures in an outline form, complete with the usual outliner functionality of collapsing subtrees, displaying to a particular depth, etc. Such a capability would support usage scenarios where the user is not certain in which category she is interested. If, for example, she was

interested in emergency room practices, she could, by browsing under the node for medical personnel, find the nodes for nurses and lab technicians. Once a node has been found, then all of the workproducts that are sorted under that node (e.g., all the workproducts for which nurses are the intended audience) can be accessed directly.

6.3.3 An Automated Scenario

All of these capabilities can best be illustrated through a sample scenario that takes advantage of all the automated capabilities discussed above. In particular, the keyword search capabilities of the web can be combined with the hierarchical structured indices to exert fine control over the dossier.

> Example. Suppose that a system developer is interested in how the duties of a midwife are carried out in a modern hospital. He is interested in knowing which personnel are responsible for which duties, and what is the sequence of back-up support in the event of complications.
>
> He begins by using the keyword filter on words such as "obstetrics," "childbirth," and "maternity." This filters the dossier down considerably, but there are still a large number of workproducts available. He begins by looking in the knowledge source hierarchical index under the Medical node in Exhibit 23 and finds that there are still too many workproducts to examine sequentially. So he descends toward Physician, and finally to Surgeon, where he finds two workproducts, one a task analysis of the Caesarean section procedure, and the other a concept diagram of drugs for cervical dilation. A search under Nurse instead again shows too many workproducts; however, a further keyword filter on "Caesarean" brings up a task analysis for preparing the patient for the procedure, as well as a comment on "breech delivery". A further search on breech delivery under Nurse shows task analyses for procedures to turn breech babies before delivery.

Navigation through the dossier can continue in this way, alternately using keywords and hierarchical structures to limit the number of workproducts accessed. If all capabilities are available on the WWW, then this capacity is accessible to members of the target audience all over the world. The WWW also provides the capability to view the KA workproducts themselves once the search has been sufficiently limited.

6.3.4 The Reuse Library Framework

Although all the capabilities described above are available in one form or another, there does not yet exist a comprehensive tool for supporting knowledge acquisition in Canvas. One candidate that provides much of the necessary support is the STARS Reuse Library Framework (RLF) [53]. RLF is a framework for creating hierarchical indices of libraries of complex objects. It was designed, as its name suggests, as an infrastructure for managing libraries of reusable software components. RLF brings with it a semantics for hierarchical models of classes and instances; Appendix B gives details of the hierarchical modeling language behind RLF, as well as examples of modeling the knowledge acquisition process itself in RLF. For the purposes of Canvas, RLF classes correspond to the nodes in the hierarchies given in Section 6.1. The instances of the classes correspond to the KA workproducts themselves.

The OpenRLF [50] is a set of WWW-based tools for browsing RLF models. The OpenRLF browser has more capabilities than we will describe here, but in order to understand its application to SEPWeb, we will outline a few of them. The OpenRLF browser presents a hierarchy in the form of an outline, and provides familiar expand/collapse display operations on the outline. In addition, the browser represents relationships between concepts as hyperlinks. Finally, the OpenRLF browser organizes arbitrary URLs by listing them under particular nodes in the hierarchy.

Dr. Jim Solderitsch at WPL Labs has developed a prototype a dossier index for SEPWeb using the OpenRLF browser. Since RLF only allows a single hierarchy, the multiple hierarchy structure of the dossier index is represented as a single, large hierarchy, which contains all the index hierarchies as sub-trees. The entries that are indexed in this dossier are just the entries publicly available in the SEPWeb.

An OpenRLF Browser demo can be seen on the WWW at:

http://rlf.wpllabs.com/rlf-docs/rlf-browser.html.

The index described here can be viewed by selecting the model name "SEPWEB Dossier" from the demonstration models found there. (This demonstration model was created as part of the same SEP/ODM method integration project which produced this guidebook.)

The outlining capabilities of the OpenRLF browser complement the keyword search capabilities already provided by the SEPWeb itself. The keywords allow a document to be accessed according to any word that appears anywhere in it, thereby allowing very flexible access. The keyword search, however, does not give a sense of the scope of the materials in the collection, or how they are related. The OpenRLF index, on the other hand, restricts the ways in which a workproduct can be looked up (since it is a single hierarchy, it restricts this even more than the multiple hierarchy outlined here), but provides a much more complete picture of whole collection, and indeed, of different subsets of the collection, and how they are related.

OpenRLF already has automated support for many of the capabilities in the scenario above, including an outliner that works on the WWW. This means that a user can examine the structure of the dossier index, and view a page from the dossier itself as needed. Further work on OpenRLF and related technologies includes support for model types other than hierarchies, and a wider array of ways to view and navigate through multiple hierarchies.

7.0 Conclusions

In this guidebook, we have presented an approach to organizing a knowledge acquisition project that includes insights gained from the point of view of the Organization Domain Modeling (ODM) and Scenario-based Engineering Process (SEP) methods. Many of the insights were gained from review of a very large knowledge acquisition effort, TCIMS, which has been a source of examples throughout the guidebook.

The guidebook has described the kinds of activities that can usefully be considered as knowledge acquisition, as distinguished from many closely related forms of learning and knowledge transfer. A conceptual framework has been offered that clarifies the role of various communities of practice (focus, investigator and target communities) and basic elements of the knowledge acquisition process (i.e., the role of investigators, informants, and artifacts in acquiring knowledge about given topics within specific settings). These are integrated in an overall metaphor of the knowledge acquisition "canvas" that includes multiple threads of interactions among the various elements. The guidebook has also presented specific recommendations for planning and managing a knowledge acquisition enterprise in light of these basic concepts, as well as implications for automated support of aspects of the knowledge acquisition process.

Throughout this guidebook, certain key principles have emerged as recurring themes in both the concepts and specific guidelines offered. This section presents our conclusions by describing some of these key Canvas principles and their implications for extending the applicability of KA, and proposing some areas for future research.

7.1 Canvas Key Principles

This section summarizes and recapitulates key principles that describe what it means to do knowledge acquisition "according to Canvas," and links them to their implications in the planning of a KA effort. These principles reflect our view of "best practice" in the KA field.

- Cultural differences between communities of practice are the greatest barrier to knowledge transmission.

  In order to make sense of a complex world, people rely on simplifying assumptions that are often difficult to articulate. Cultures grow around shared sets of these assumptions. Some of the techniques for knowledge acquisition explored in this guidebook are adapted from social science research, where "culture" in this sense may be interpreted quite broadly; but the techniques can also be adapted to the "micro-cultures" of particular organizations, technology development environments and work settings.

  Participants in a knowledge acquisition project play an "ambassador" role in making cultural assumptions explicit so that information can more readily flow across the boundaries of different communities. Awareness of the distinct communities involved, and care in mapping terminology, etc. from one community to another, are key aspects of a systematic approach to knowledge acquisition.

- Anyone can be a source of knowledge.

  So-called "experts" are not the only people who have knowledge; everyone who accomplishes some task exhibits knowledge. No community or practitioner has a monopoly on "important" knowledge. A comprehensive picture of an interaction among communities of practice will include information from all of the communities involved, and a broad selection of practitioners in each of those communities.

- Bias is unavoidable, but its effects can be managed.

  As human beings try to make sense of their world, they take advantage of patterns found in their experience. While this makes the flood of information in the world intelligible, it also means that one's interpretation of the world depends on one's own background. Through systematic tracking of exposure to information, bias in a knowledge acquisition project can be controlled. This requires careful planning and record keeping of the activities of each actor in the knowledge acquisition team.

- Variability is an integral part of knowledge.

  People will disagree on almost any topic. Some disagreements are the result of one person being misinformed, but the most interesting disagreement comes from difference in background, interests, or viewpoint.

  It is easy to adopt strategies that minimize variability in KA, but in so doing we may lose a rich source of data. Differences need not be ironed over in collecting a repository of knowledge; the differences themselves are part of the rich, cultural web of the knowledge. For some purposes, such as domain engineering, variability is an essential aspect of the data to be gathered. A number of representational notations can express this variability, leaving the audience to decide which variant is appropriate to any situation.

- The process of eliciting knowledge intervenes in the studied setting.

  The knowledge acquisition process will cause people to think about their work in a different way, and to make new connections to other workers. The mere act of reflecting on their work or meeting with other people will cause change in the focus community. The knowledge acquisition practitioner can use this insight as a way to create maximum value from the acquisition process itself. Intervention, properly controlled, can be an instrument for improving work practice.

- Representation is key to managing knowledge acquisition.

  The knowledge acquisition process is not complete until the knowledge is written down. This process of distillation and translation can present great challenges for the investigator. Representations should be chosen carefully based on the communities where they will be used, and by the individuals who will verify their content. Representations have inherent bias. Some can accommodate variability, while others cannot. Representations are a record that can have value in the studied setting.

For some readers, these principles may seem to be obvious statements of common sense, too plain to merit mention. Experience demonstrates that considerable skill and care is required to apply these principles systematically on a knowledge acquisition project. Some readers may be daunted by the amount of effort required to systematically consider these principles in KA. We believe that simple awareness of some of the principles can improve practice; and that more systematic application will be essential to any KA effort of significant size. To other readers, these principles might challenge assumptions they consider incontrovertible or self-evident. We believe they are essential principles to acknowledge in order to make consistent and dependable use of knowledge acquisition results, and to better understand the potential value of those results.

7.2 Future Research

We conclude by giving a few ideas of directions for further research that would aid the knowledge acquisition process. Some other ideas along these lines, particularly focused on supporting technology development, have already been discussed in Section 6.3.

7.2.1 Presenting Knowledge to Various Audiences

Given the central role that representations play in the knowledge acquisition process, much future work that we propose in this area focuses on presentation and translation of representations. Representations of knowledge in Canvas are typically produced by an investigator, but they must be read by other people, including the informants from which the knowledge is elicited and the target audience for whom the knowledge is intended. This means that there is a need for careful study of how representations can be presented to various audiences in such a way that the audiences will not misconstrue the representations' content.

A clear example of this problem is the situation with taxonomies mentioned in Section 5.6.4; many communities use taxonomic-style diagrams for many purposes. If we want to present a taxonomy to someone in one of these communities, we need to make certain that we do not collide with any use of similar diagrams with which practitioners are already familiar. This requires a careful study of how people understand representations, including analysis of psychological as well as cultural aspects. Furthermore, some notations have a tight semantics, for which technical training is needed in order to understand the semantics. Some communities might be willing to make the investment to undergo this training (in much the way that psychologists have adopted the study of statistics into their practice), while others might not. At the very least, notations with clear semantic descriptions, as well as automated support, could form the basis for such adoption by a community of practice.

Another issue when presenting a representation to some audience is the cognitive load that the representation places on its viewer. Again using taxonomies as an example, simple trees can be presented a few dozen nodes at a time using presentations such as outliners and tree browsers. Interactive support for such structures allows a user to manage several dozen such objects. On the other hand, if a tree includes more detailed structure (e.g., semantically laden linkages between the nodes, or attributes of the nodes), then the basic outliner and tree browsers lose their effectiveness. The RLF diagrams shown in Appendix B demonstrate the difficulties of showing several interconnected nodes in a semantic net, and a simple solution that was effective in this case.

Future research on these issues would include study of how practitioners from different communities understand various representations. The research could also include experimental work in the spirit of psychological investigation about how well subjects can cope with information, as well as cultural studies of ways in which similar representations are understood in different communities.

7.2.2 Translation Between Representations

One of the constraints on representations in a knowledge acquisition context is that they typically have to serve two masters; first, they have to be understandable to someone who is in a position to verify the correctness of the knowledge represented, and second, they have to be understandable to someone in the target community, which is often more technically oriented. This problem cannot simply be ignored by training one community to understand the notations of the other; many notations are quite technical and require expertise to use properly. Because of the possibility for misunderstanding, poorly formed representations can be more dangerous to cross-community communication than having no representations at all.

One solution to this problem is to let the investigator community play the bridging role by performing translations by hand. The investigators can work with knowledge sources in representations that are familiar to the source community, verifying the information represented. Then they can translate this information into a form that is understandable by the target community. It is then up to the investigators to make certain that the information is faithfully transferred from one representation to another.

Tools could facilitate this endeavor. If all notations involved have a formal semantics, it would be possible to prove whether one representation has the same content as the old. The translation could even be automated; given the semantics of both notations, representations in the source notation could be used to automatically generate representations in the target notation. A problem with this approach is that often the power of a representation lies in the informality of its notation, which leaves certain things unspecified or ambiguous. Translations between informal notations, or from an informal to a formal notation are problematic. Providing tools to support such translations in the context of knowledge acquisition would be a great boon to project planning.

In some collaborative settings, the investigator provides the tools for representing the informant's knowledge in the form of a computer program. In this case, the expert constructs his own models with the help of the program. The Repertory Grid method [36] provides extensive model visualization tools, allowing the expert to input and manipulate his own models. An extreme version of this in the context of expert system construction is reported by Welbank [58], in which the expert learns to program a rule-based system, and maintains the rule models of the expertise himself. This can raise problems in practice if the size of the rule base grows beyond the expert's patience to maintain it, but there is no barrier in principle to having an expert learn how to construct even very sophisticated models. Our long-term goal is supporting technology that allows a flexible choice of representations for KA, ability to transform models among these representations, and a support for many different collaborative modes of knowledge capture, exchange, and creation.

7.3 Final Thoughts

A primary reference point for Canvas is the role of KA in building better software systems. What level of maturity is needed to do this well? We believe that collaborative KA is an appropriate paradigm for this kind of KA situation. When software developers and end-users are both viewed as potential informants, the assumption that informants cannot themselves do reflective, taxonomic or knowledge-creating activity breaks down. Software technology artifacts (systems and the other workproducts used in creating them) lose their status as "received wisdom" and can be viewed as cultural artifacts, reflecting embedded views and assumptions and amenable to revision.

We have tried to demonstrate that knowledge acquisition involves a coordinated set of cognitive and interpersonal skills that are as valid an area of specialty as technology expertise. Knowledge acquisition is, in some sense, the fundamental intellectual activity of a social creature, gaining knowledge from its fellows. The process of acquiring knowledge has an intrinsic value that usually goes beyond the value initially anticipated.

Availability of such expertise can add value to a technology development project, but may need to be managed in distinct ways. There is nothing inherent in the training of technologists that make them particularly capable at this set of skills. Thus "drafting" technologists to play this role without special training or acknowledgment of the separate processes required may not yield the benefits of systematic KA. The requisite skills also challenge some assumptions of traditional knowledge acquisition researchers. Ideally, the Canvas investigator would approach KA as an opportunity to facilitate knowledge creation and sharing between technologists' and end-users' communities, potentially transforming the technology development process itself as a result.

Appendix A: Canvas as ODM Supporting Method

Although Canvas does not assume a domain engineering context, it directly supports knowledge acquisition as part of the ODM domain engineering life cycle. In theory, Canvas could also be applied to other domain engineering methods and approaches. However, one distinction of ODM is its modular structure: it provides a core process model that can be integrated with a set of supporting methods, of which knowledge acquisition techniques are one of the most central areas required. A Canvas-based approach to KA planning can therefore be readily integrated into an ODM project.

In addition, many core concepts incorporated in Canvas are directly transferred from their use in ODM, so the approaches are particularly compatible at both the conceptual and process levels. Although Canvas focuses on KA planning aspects more than detailed elicitation techniques (e.g., interviewing protocols), this planning framework is the most important element needed to incorporate such techniques into the ODM life cycle.

This appendix discusses Canvas as an ODM supporting method. Section A.1 describes the major common concepts that facilitate an integrated ODM/Canvas approach to KA for domain engineering. Section A.2 gives an overview of the ODM domain engineering life cycle, and discusses the specific interfaces between the ODM process and the planning process encompassed by Canvas. Section A.3 presents some further guidelines to consider in carrying out Canvas-style KA planning within the ODM context. The material does not assume familiarity with ODM, but also cannot provide a thorough introduction here. Readers are referred to the extensive documentation on the method in [49].

A.1 ODM and Canvas: Common Concepts

A number of the concepts introduced in this guidebook derive from or are in accord with the ODM perspective. These include emphasis on the following points:

- Grounding the project in an explicit analysis of the various stakeholder issues involved, the "organization context";
- Formal definition and scoping of the domain as part of planning
- Context recovery as an essential process in utilizing system artifacts as data, and allowing for the effects of bias;
- Systematic approach to gathering data with respect to ranges of variability in work practice and system function in different settings; and
- Recognition that domain engineering (and KA) are interventions in the organizational systems in which they are performed.

Stakeholders

ODM reflects the idea that domains are socially defined agreements about an "intended scope of applicability." Domains are always grounded in some "organization context," a community of interest that can take many different forms, such as a company or division, a consortium of multiple organizations or a standards organization. Stakeholder issues, always a potential problem in any project, are critical risk factors in domain engineering, which by definition involves designing for multiple contexts of use.

ODM begins with systematic exploration of stakeholder interests to guide selection, scoping and definition of domains strategically aligned with the business interests for the organization(s). This stakeholder context provides further grounding throughout the subsequent phases of the process. The domain engineer examines a wide network of stakeholders and situates his or her own interests within that network. This is important for insightful and accurate data analysis and interpretation that captures how the players and their perspectives are situated socially by their context. Canvas extends this stakeholder analysis in more detail to the knowledge acquisition phase, considering the stakeholder interests informants in the focus community and the shifting stakes of the domain engineering team as they take on a KA investigator role. Canvas specifically draws its informants from various stakeholder groups, and different kinds of practitioners in those groups.

Domains

One of ODM's primary contributions as a domain engineering method has been focus on the process and method challenges in defining and scoping ***domains***. In initiating a Canvas-based KA enterprise in the ODM context, this domain definition would provide the initial organizing basis for selecting the topics of interest. The ODM notion of domain could be applied in the Canvas setting to set clearer boundaries around and distinction between topics, even outside of a formal domain engineering context. It is important to note that, in ODM terms, the very process of defining the domain can be an organizational intervention, since the domain boundaries may or may not be familiar to practitioners as a conventional way of dividing up the functional concerns of the system(s) under investigation. When KA is to be performed for such a narrow domain, particular care must be taken in filtering the data and working with the conceptual terms of the informants.

When using KA as a kind of pre-requirements analysis for system building, there is a danger of lack of focus. To some extent all work practice taking place in the examined settings may appear to be relevant for knowledge acquisition. Though technologists intending to build systems are the primary audience for the KA materials, it may be difficult to know how to filter the materials acquired to give the technologists only the relevant data. There may be an indistinct notion of what technology is intended to be introduced. There may be only weak correlations between technology specifications and the work practices that must be studied in order to assess the viability of the technology. Particularly for innovative technology, it may be very difficult to get end users to provide data about system capabilities which they have not yet used or even seen.

The implication is that technology development goals form poor mechanisms for bounding and scoping knowledge acquisition. Investigators are likely to set themselves the goal of describing complete scenarios of work practice within the settings, with the aim of getting the data first and deciding on relevance later.

Context Recovery

One of the primary purposes of domain modeling according to ODM is *context recovery*, which is the process of identifying and making explicit the assumptions embedded within software artifacts. These assumptions come from the cultural context in which the artifacts are developed and used. This means that domain modeling in ODM bears considerable similarity to cross-cultural investigation of the sort found in ethnographic studies. However, because of the particular constraints of domain modeling for software, knowledge acquisition for ODM imposes some general requirements that go beyond standard practice and general principles and skills of ethnography.

ODM shares with traditional ethnography an attention to the effects of bias on the information gathered from the culture to be studied. On the other hand, the ODM approach recognizes a wider range of stakeholders than many other ethnographic endeavors. Whereas any ethnographer tries to impose little on the culture and to be aware of the politics of "us" and "them," little work has been

done in ethnographic research equivalent to what the management and consulting world would call alliance forming and management, consultant-client contracting, or stakeholder analysis. Canvas provides support for the traditional ethnographic needs of ODM by recognizing knowledge acquisition as a process of communicating between work cultures. Canvas also provides for more than a simple two-player ("us and them" or developer/user) model of stakeholder groups and communities of practice involved in KA. This could be applied more broadly to the entire domain engineering life cycle thus enabling a consistent and integrated approach.

Variability

The ODM approach is centered much more on the study of variability than is common in ethnographic studies. Since the result of a domain modeling project according to ODM is an asset base that can be used by many stakeholders in many contexts, the variability inherent in the domain must be elicited, maintained and managed throughout the ODM lifecycle. Canvas specifically provides support for eliciting and representing variability.

Intervention

ODM views the overall domain engineering life cycle as an intervention in the various stakeholder organizations involved, and includes within its knowledge acquisition guidelines a cycle of elicitation and representation that accounts for iterative and ongoing feedback of information to the informant. The Canvas model of knowledge acquisition as intervention fits well into this pattern, by recognizing the effects that the process of knowledge acquisition itself can have an impact on the focus community.

A.2 Interface between ODM and Canvas

The core ODM process is structured into three main phases: *Plan Domain*, *Model Domain*, and *Engineer Asset Base*. ODM phases are iteratively decomposed into sub-phases and tasks; each task in turn is documented via sequences of or alternative activities and a structured set of work products that integrate information gathered throughout the entire process. Major results of each phase—domain definition, domain model, and asset base—are targeted to provide direct benefit to the project stakeholders, as well as to provide a systematic starting point for the next phase.

Information Acquisition techniques are a major supporting method area within ODM, which fit into the overall ODM process within the sub-phase called *Acquire Domain Information*. Information is gathered throughout the process; but within this sub-phase systematic data-gathering becomes essential. Exhibit 26 shows the task structure of the ODM process, and where the *Acquire Domain Information* sub-phase is located.

The scope of Canvas is broad enough to allow its use as a supporting method of ODM for the *Plan Data Acquisition* task within the *Acquire Domain Information* sub-phase. Within the core ODM method, the *Plan Data Acquisition* task involves several key activities:

- Selecting a representative set of systems to study to produce the domain model;
- Selecting the settings that will be studied for each representative system;
- Selecting the information sources that will be consulted for each representative system and setting. Consultation of sources could include interviewing human informants, observing domain-related processes, or analyzing artifacts;
- Characterizing biases of investigators.

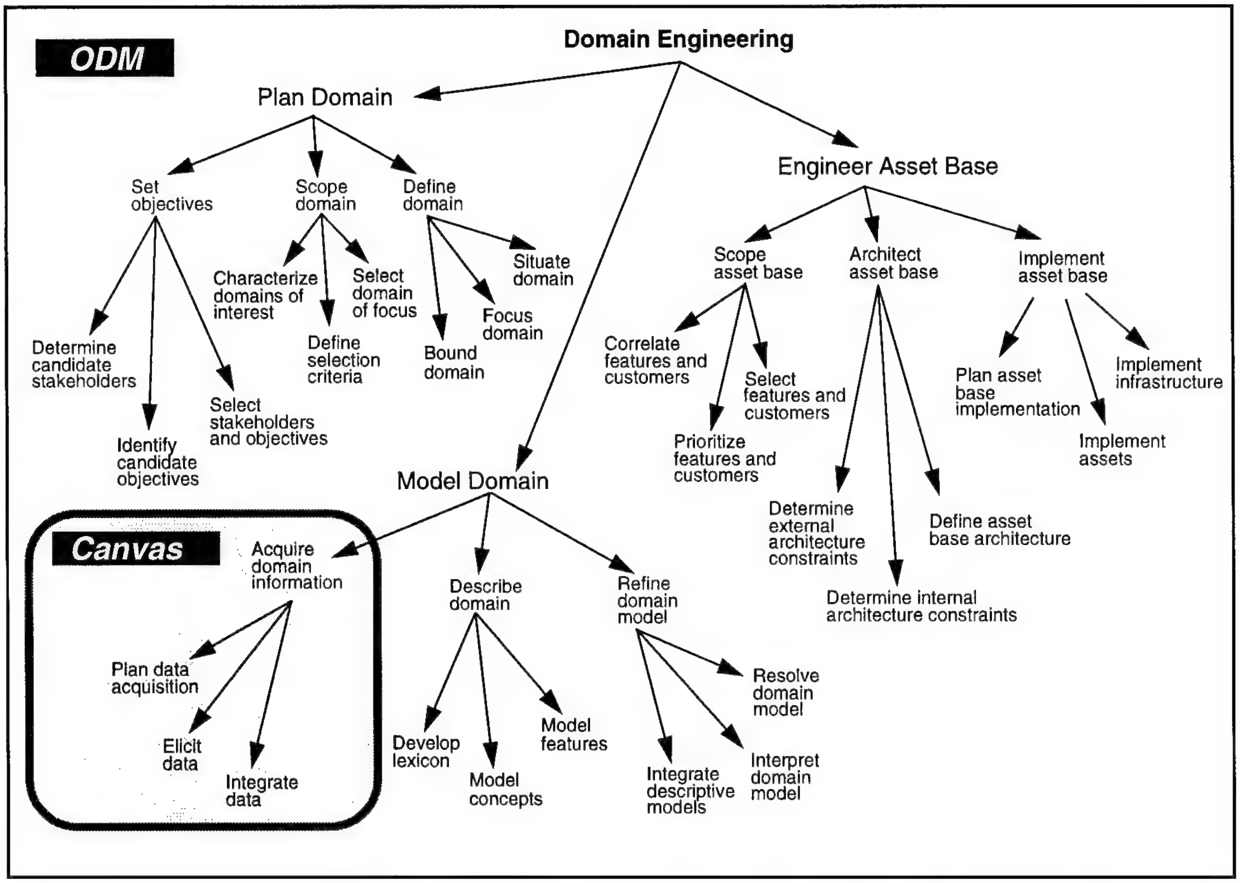


Exhibit 26. Canvas as a Supporting Method of ODM

Canvas offers guidance to support each of these activities, though it does not provide a formal process model for the KA planning process at the level of the ODM process model. This sub-section will describes how to integrate a Canvas approach as a supporting method for the ODM *Plan Data Acquisition* task.

Knowledge acquisition goals (required as part of the planning process) will be shaped in large part by the fact that the KA enterprise supports a domain engineering project. This means, for example, that management and documentation of variability will be an essential part of the process. It also means that, for domain engineering projects in software-intensive domains, KA techniques specifically addressing technology-intensive settings will be required.

Canvas provides several points at which knowledge sources to be consulted are determined (Section 4.5). These selections are based primarily on properties of the information source, including issues of currency and accessibility. In ODM, there is another chance to scope the set of knowledge sources to be studied. As part of the *Plan Domain* phase of ODM, stakeholder issues are used to determine what belongs in the domain and hence, what is a candidate for being selected for study during knowledge acquisition. This means that when performing Canvas as part of an ODM project, there is more guidance for deciding what will be studied and what will not.

When invoked in the ODM context, Canvas KA planning has a number of ODM workproducts and data items available as controls and inputs that can support the planning process. These are described below. A full description of these workproducts and data items, their content and devel-

opment can be found in the ODM Guidebook V2.0 [49]. In this section, we will follow the conventions of that guidebook by printing workproducts in SMALL CAPITALS. Relevant points of interest are discussed in the context of each potential link between the overall ODM project and the KA enterprise planning process.

ODM Controls to Canvas

The following workproducts and data items constitute controls on the *Plan Data Acquisition* task[1]:

- PROJECT OBJECTIVES. In addition to the general KA goals established due to the domain engineering nature of the project, the specific objectives of the project will have direct impact on the KA enterprise goals. For example, a domain engineering project may be initiated to capture veteran expertise in application development. In this case, the dossier created by Canvas (called the DOMAIN DOSSIER in ODM) will be of direct use to the organization as codified documentation of expertise. If PROJECT OBJECTIVES include the discovery of innovative features that might provide competitive advantage to a software product line, then direct observation of end-users interacting with legacy systems might take place, to elicit opportunities for new features.

- PROJECT CONSTRAINTS. These constraints can include budget, schedule, access constraints as well as constraints such as use of proprietary data.

  Up-front acknowledgment of and planning within constraints is part of any systematic KA process. In general, it is very easy to spend too much time gathering data. Without a budget and plan that shows how much information can actually be used, it is easy to expend scarce resources collecting data that will not be able to be effectively utilized in modeling. Similarly, it is easy to base an initial plan on the assumption that access to some key informants will be forthcoming, only to find that such access is extremely difficult to obtain in reality.

  Constraints are particularly relevant in the domain engineering context. Data is generally being collected from multiple exemplars, which means that the focus of data collection and level of detail become essential controls, since an error in gathering too much data may be repeated on multiple systems.

- DOMAIN DEFINITION. In ODM domain engineering, the DOMAIN DEFINITION provides a tight focus for data acquisition, circumscribing both a specific set of exemplar systems that are included in the domain, and a specific focus area within the overall functionality of systems in the domain. This is one advantage of using Canvas in the ODM context as opposed to within a general system engineering context.

  This does not mean that in a given interview only domain-relevant data will be collected. Informants may not have access to or understanding of the concept of the domain boundary (as refined in the earlier *Define Domain* sub-phase of ODM) as a basis for discourse or focusing attention. Information must be acquired in chunks and sequences that the informants find workable, and subsequently filtered. Yet the DOMAIN DEFINITION can still help investigators to steer inquiry in more focused directions and to rapidly filter the data acquired to extract domain-relevant material, while the session is still comparatively fresh in their experience.

---

[1]. The term *controls* here is a technical term from the $IDEF_0$ process definition method used in documenting ODM, and described in and [41]; the STARS approach to process definition is described in [21].

ODM Inputs to Canvas

The following ODM workproducts and data items serve as inputs to the KA planning processes addressed in Canvas:

- DOMAIN STAKEHOLDER MODEL, STAKEHOLDER DOSSIER

  The DOMAIN STAKEHOLDER MODEL and the STAKEHOLDER DOSSIER contain information about the stakeholders for the domain engineering project. The dossier includes all information gathered about stakeholders and their positions, and can include such items as company prospectus, commercialization plan, interview notes with key personnel, etc. The model includes the relationships among the stakeholders and their roles in the project. This information is not only needed to determine what information will be gathered; it is also a primary source of information for finding informants willing and able to participate in the knowledge acquisition project.

- DOMAIN STAKEHOLDER KNOWLEDGE

  This data item refers to all information that is available from stakeholders. For the purposes of knowledge acquisition, it is the stakeholders who are chosen as informants who are of interest; the domain stakeholder knowledge then refers to informant knowledge. Access to this information is gained in Canvas mainly through interviews.

- EXEMPLAR SYSTEM ARTIFACTS, DOMAIN ARTIFACTS

  The other main source of information in Canvas is artifacts that are examined. ODM divides artifacts into two categories. EXEMPLAR SYSTEM ARTIFACTS are documents, notes, code fragments, executable systems, etc. pertaining to a single software system to be studied. DOMAIN ARTIFACTS are any materials pertaining to the domain, that are not related to one specific system. Examples of such artifacts are survey articles, comparative "consumer's reports" style charts, or even models created by previous domain engineering efforts. The Canvas framework allows for both of these types of information.

Canvas Outputs to ODM

The following Canvas results contribute to ODM workproducts derived from the *Acquire domain information* sub-phase:

- KA enterprise objectives called out in this document in Section 4.3.3 correspond closely to the "data acquisition goals" created as part of the DATA ACQUISITION PLAN in ODM. These should be aligned closely to the overall PROJECT OBJECTIVES.

- The list of elements selected in Section 4.5 constitute the REPRESENTATIVE SYSTEMS SELECTION in ODM. The criteria used in Canvas for selecting these representative systems draws upon information available in the STAKEHOLDER DOSSIER and the DOMAIN STAKEHOLDER MODEL. What constitutes a "representative system" in the ODM context is typically a single software application; but the relevant settings to study for that system could include both the developers' and users' communities. So selection of communities of practice, settings, artifacts and informants will, in the ODM context, be constrained by selecting the overall set of representative systems to study.

- The structure for the dossier that is determine in Section 4.7 is the "bootstrap" for the DOMAIN DOSSIER that is built and maintained throughout the *Acquire domain information* sub-phase in ODM.

A.3 Guidelines for Integrating Canvas and ODM

Although Canvas fits very well with a domain engineering project using ODM, Canvas does not assume application in a domain engineering context. When we know that Canvas is being used in an ODM context, we can be more specific about guidelines for certain activities. Other activities can be de-emphasized, because the ODM context provides some information that might not otherwise be available. On some occasions, both ODM and Canvas cover similar territory, and the combination allows for choice of whether a particular decision should be made as part of Canvas or as part of ODM.

When Canvas is performed in an ODM setting, the subsequent ODM *Model Domain* sub-phase creates a model from the knowledge in the dossier. This means certain modeling activities need not be performed as part of knowledge acquisition, but can be left for the domain modeling stage of ODM. However, depending on the choice of representational notation, the initial representation and codification of knowledge in Canvas could be similar to domain modeling process as understood in ODM.

In these cases, there are two complementary ways to view the interaction of the two processes:

1) The entire ODM *Model Domain* sub-phase (encompassing planning, data gathering and integration of data, modeling and interpretation) can take place in the course of collaborative knowledge creation with informants in the focus community.

2) Alternatively, from the Canvas perspective, the team who will be doing the domain modeling (as distinct from the KA or investigator staff) can be viewed as the immediate audience or target community for the knowledge acquired. This team might be the same team who is doing the knowledge acquisition, or might be another group. In any case, the set of representational notations with which the team is familiar will include some modeling notations, and it can be assumed that the audience is "model savvy".

The level at which planners view the unfolding of the knowledge gathering, interpretation and integration cycle, or the broader ODM cycle of descriptive modeling, interpretive and innovative modeling, must be determined uniquely for each project as part of the planning process. Similarly, where knowledge acquisition leaves off and modeling or implementation of reusable assets begins is not a hard and fast line. In fact, since the dossier may be used directly by the focus community as a source for codified expertise, it can be viewed as an asset in its own right. In that view the Canvas and ODM processes converge much more closely.

Representation versus modeling

In Canvas, the representation of knowledge plays a key role in the process, since it is through representation that knowledge is made available for verification by the focus community, or transferred to the target community. The extent to which modeling must be done depends on the characteristics and requirements of the target community; the target community might require complete, semantically sound models in the dossier.

Variability management

Closely related to the previous point is the handling of variability in ODM vs. Canvas. The act of making comparisons and determining just what parts of the knowledge represent commonality and what parts represent variability is a demanding task. Since the Canvas approach stresses that variability be handled as part of the knowledge acquisition effort, when Canvas is performed on its own, an explicit representation of variability will have to be included in the dossier.

When Canvas is performed as part of an ODM domain modeling project, there is a systematic process in which variability will be modeled. This does not mean that KA practitioners using Canvas in an ODM setting can just forget about variability! Treatment of variability during knowledge acquisition in Canvas includes realizing that the knowledge represented need not conform to some common agreement among diverse sources. When Canvas is performed in an ODM context, the varying viewpoints must still be collected; only detailed modeling of which information represents differences and which represents commonalities can be delayed until the later ODM *Describe domain* sub-phase.

Appendix B: Representing the KA Process

The production of Canvas from information that was previously only available to the ODM and SEP teams individually involved a process of knowledge acquisition itself. Along the way, we used our own tools to help us to keep track of the material we were collecting. In particular, we used RLF, a hierarchical modeling notation often used in conjunction with ODM, to represent the relationship among the terms in the lexicon and in the definition of knowledge acquisition according to Canvas. We also produced an interaction diagram, similar to the ones common in SEP, to describe how the various elements interact.

This section presents the fundamental concepts used in Canvas, modeled primarily using a subset of KNET [12] representations. KNET provided the basis for the underlying semantic network representation of the STARS Reuse Library Framework (RLF) ([53], [54]) which was subsequently evolved into the OpenRLF system (as described in [50]). The OpenRLF toolset was used for the example index into the SEPWeb on-line dossier described in Section 6.0. However, the examples in this section do not correspond to the implemented indices: the graphical notation here is not equivalent to that used in the tool interface; not all capabilities implied by the modeling approach taken here are incorporated in the existing tool.

This section can be used in conjunction with Appendix C: "Canvas Lexicon" to gain an understanding of how these concepts are used in Canvas. The hierarchical aspects of KNET have been used to show the inclusion (is-a) relationships of the various terms to one another, while the relationships of KNET have been used to show other relationships. In this appendix, we use a small subset of the full features of KNET. However, the models provided here are still example fragments rather than a complete formal semantics for the terminology used in Canvas, a task well beyond our scope here. In fact, an important point of the examples is that there are alternative, legitimate ways of modeling the same data; this will be expanded on in Section B.3.

B.1 Fundamentals of KNET

KNET was derived from knowledge representation languages like KL-ONE. For readers already familiar with KL-ONE style languages, a few of the features of KNET may not be familiar; however, these are more advanced semantic features which we will not use in the examples in this section. For readers unfamiliar with these types of languages, the following description can be used as a brief primer to understand this appendix.

In KNET, the principal entities are categories, relationships and instances or individuals. A category models a class of things, such as the class of all learning interactions or the class of all Knowledge Acquisition Sessions. A relationship describes the structure and properties of categories. This appendix does not make use of KNET instances.

Categories in KNET are organized into a specialization hierarchy. A category C1 specializes another category C if C1 represents a subset of the class described by C. This means that the most general category appears at the top of the hierarchy with more specific categories below it and the most specific categories at the very bottom. For example, in Canvas, Interaction and Study are two kinds of Knowledge Acquisition Session. In the diagrams within this appendix, specialization is shown by a straight bold arrow from the specific category to the more general.

Relationships in KNET describe the structure and properties of categories. For example, a Knowledge Acquisition Session is performed by an Investigator and produces a Workproduct. Such qualities are represented in KNET by associating relationships with a category. For example, Knowledge Acquisition Session includes relationships for "performed_by" and "produces". In

this appendix, relationships are shown as a thin broken arrow from one category to another. The break is annotated with the name of the relationship. (In standard KNET, relationships are annotated with ranges indicating maximum and minimum number of instances that can fill the relationship; these are not used in this appendix.)

Relationship differentiation allows a relationship to be described in a more detailed way than is possible with a single relationship. Differentiation can thought of as specialization of relationships and is denoted with dashed arrows between the "elbows" in two broken lines. For example, a session has as its knowledge source any number of authorities; authorities who are practitioners in the focus community are known as informants, while documents and other materials from the focus work setting used as "provenance" or a basis for authority are known as artifacts. This captures the notion that "artifact" and "informant" are stances that one takes toward workproducts and focus practitioners respectively, and that these stances are just two special cases of the stance of using any authority as a "knowledge source."

Use of the KNET Model of Knowledge Acquisition

We provide a KNET model of knowledge acquisition terms for three different reasons. It is not necessary to understand the subtleties of KNET to obtain all of these benefits.

1) *The specialization hierarchies help to clarify relations between terms.*

   The definitions of terminology given in the lexicon (Appendix C) make heavy use of other terms. Many of the relationships between terms are captured in the hierarchy structures of the KNET diagrams.We encourage the reader to flip back and forth from the lexicon to this appendix to see how the terms relate.

2) *The relationship networks serve as a summary of the structure of the KA process.*

   Many subtleties, such as the fact that a KA workproduct has a particular audience, which is a community of practice, are explained in the text. Once these explanations have been absorbed, the relationships in the KNET diagrams are a concise way to record and present this information.

3) *The model can be used as a "starter set" for a dossier of KA workproducts.*

   Any aspect of the planning process could be used to index the dossier. The most likely parts of the KNET model to be used as an index are the categorizations of audiences and the types of representation.

B.2 KNET Model of Knowledge Acquisition

The KNET diagram for the entire knowledge acquisition process is too complex to display in a single display; we have therefore broken it into three pieces. Each piece can be considered a view into the common model, focused on a single category. All specializations of that category are shown, as well as all relations defined for the focus category or any of its specializations (i.e., any relation arrows that lead *away* from the category of focus or its specializations). Any category that must be shown in order to display these relations is also shown. There is no guarantee that specialization hierarchies for non-focus categories are complete as shown. In order to find complete specializations for a non-focus category, it is necessary to look in the diagram where that category, or one of its generalizations, is the focus.

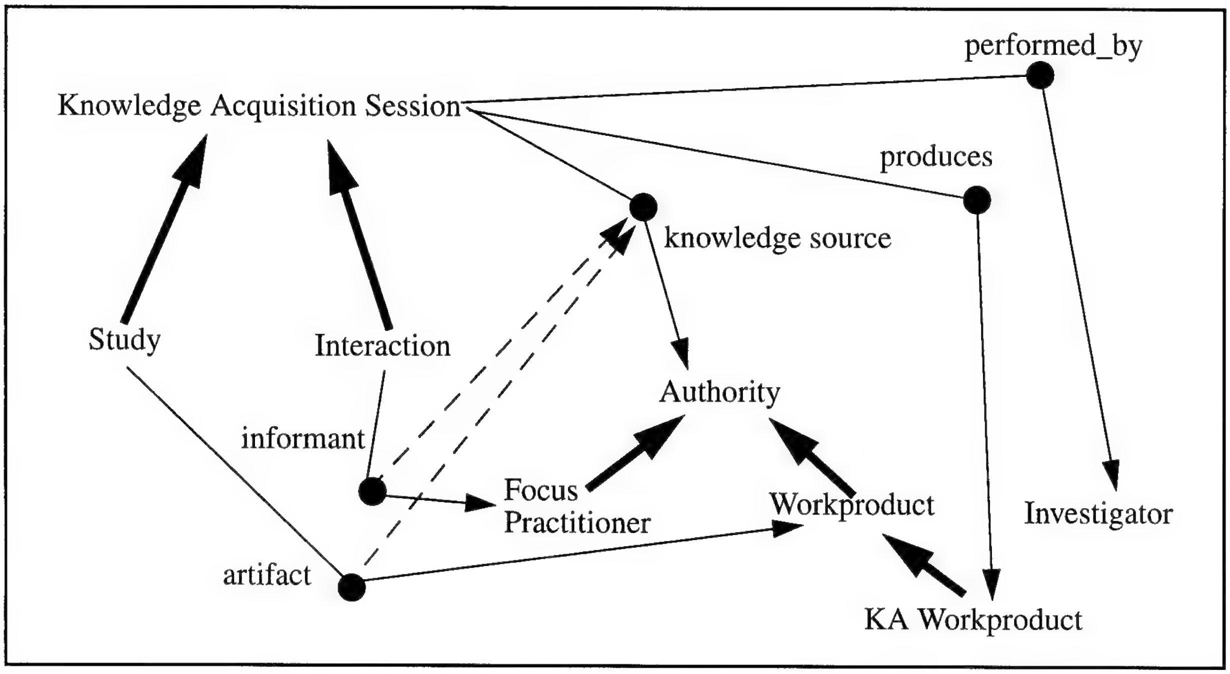


Exhibit 27. Relationships Centered around the KA Session

Exhibit 27 shows the KNET diagram for the knowledge acquisition session itself. A knowledge acquisition session is performed by an investigator, has a knowledge source that is an authority (about some topic, not shown), and produces a workproduct. In the case where the knowledge source is a human being, the session is called an interaction. When the knowledge source is a workproduct, the session is a study. Artifacts and informants are differentiated forms of knowledge sources. When describing a session, we typically refer to informants, investigators and knowledge sources, rather than focus practitioners, workproducts and authorities because we are interested in their roles in the session. That is, the status of being an "artifact" is a relationship rather than a fixed attribute of a workproduct.

Exhibit 28 shows the KNET diagram for the practitioners involved with the knowledge acquisition effort. A practitioner is a member of some community of practice. The two specializations of practitioner, focus practitioners and investigators, are members of the focus community and investigator community respectively. When the meaning is clear from the context, we often refer to focus practitioners simply as practitioners.

Exhibit 29 shows the KNET diagram for the community of practice. In Canvas, we are interested in three kinds of communities: an investigator community, a focus community, and a target community. At a lower level of detail, every workproduct has an audience that can be characterized in terms of some community, and the members of each community practice in some work settings.

B.3 Alternative Representations

KNET is an example of one style of taxonomic notation, as described in Section 5.6.4. In this appendix, we have used a subset of KNET (i.e., without cardinality or restriction of cardinality, and leaving out different types of differentiation) as a display mechanism. In order to make the diagrams legible, we have added some extra parameters, as described in Section B.2, to control the display. By using selectiveness, multiple displays, subsets of semantics, and similar techniques a tailored representation and notation was created to suit a specific intended audience.

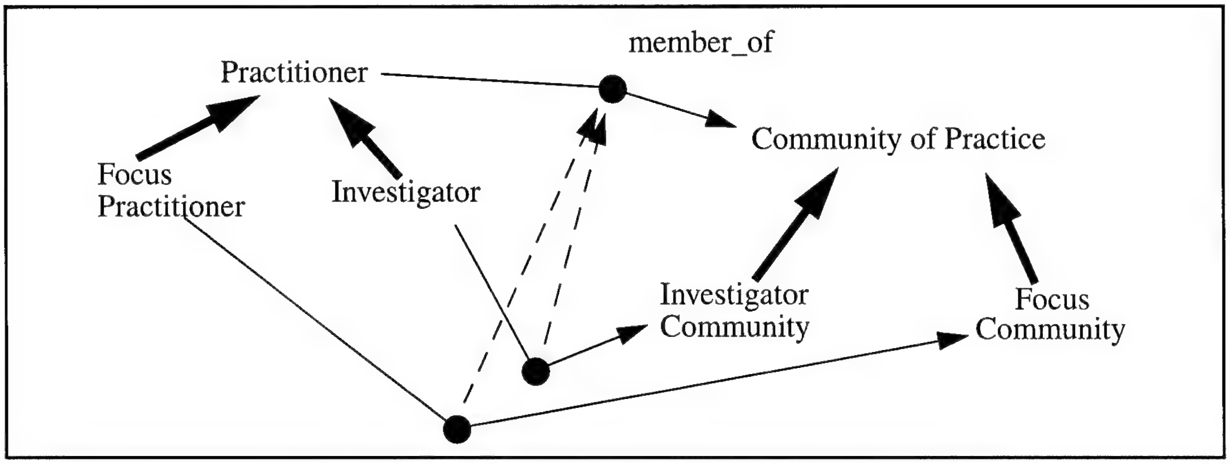


Exhibit 28. Relationships Centered around Practitioners

Exhibit 29. Relationships Centered around Community of Practice

The result is a notation that could be used for a knowledge acquisition workproduct. Here, the topic of the workproduct is the knowledge acquisition process itself. A different representation of much the same material can be found in Appendix C in another notation, that of a lexicon.

As yet another alternative view, Exhibit 30 presents a top-level view of the interactions of Canvas represented in a SEP interaction diagram, a variation on the Petri net. The primary distinction between interaction diagrams and classical petri nets is that objects (items represented in ovals) may retain their identify through transitions (rectangles). Transitions are places where objects are created, destroyed, modified, change state, or interact with one another. The iterative nature of the interactions is evident.

As represented in Exhibit 30, sessions conducted by the investigator(s) produce work products using as input artifacts from the domain and/or the information provided by interviews and/or observations involving informants from the domain. Some artifacts are carried over from prior work in ODM (i.e., legacy systems, design specifications, user manuals, and other documents.) Other artifacts are derived from prior work in SEP, such as representative scenarios, concepts, and task decompositions. Transfer is the process that models the incorporation of Canvas work products into the understanding of the domain by the audience

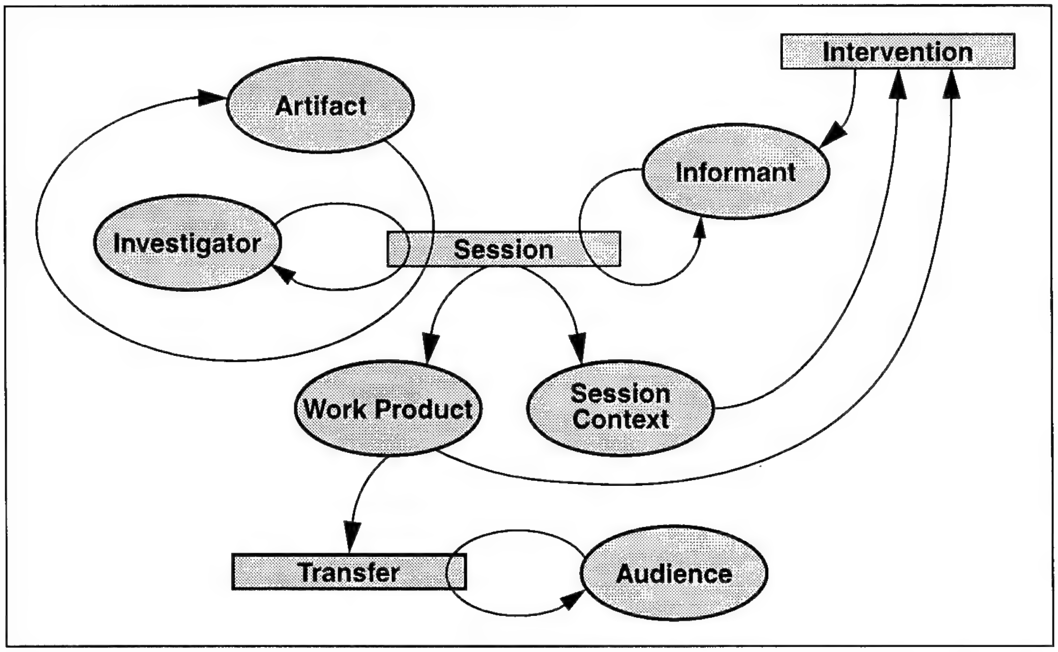


Exhibit 30. Interaction Diagram for a Knowledge Acquisition Session

The completion of the feedback loop for investigators provides for iterative improvement of the domain context held as corporate knowledge by the investigator community. The production of the KA workproduct, and the context of the setting itself, can intervene on the informant's work practice, either individually, by clarifying his own understanding of his field, or as a community, by fostering connections that had not otherwise been made.

Epilogue: Using KNET as a Knowledge Acquisition Tool

The audience of these workproducts is you, the reader of this guidebook. However, there might be other interested parties who will become an audience for the dossier. In this case, the authors of this guidebook were an audience for the diagrams. As is often the case, we began examining the diagrams as a way of validating the information we were gathering. As time went on, however, we began to use them as a reference for our own work while writing and revising the guidebook.

Each notation brings with it certain views on the information that others do not. For example, in Exhibit 27, we needed to find a name for the category of entities that could serve as knowledge sources. The two known specializations of this new category (focus practitioner and workproduct) had stories to be told about them in their own rights (as the subject of two of the threads in the Canvas.) However, the new category had not played a role in the Canvas exposition so far (and did not appear in the lexicon). The formal constraints on the KNET diagrams forced us find a name. This process revealed any number of stakeholder issues that had not been evident before (such as why "knowledge source" had to be a relation, not a category; which category the name practitioner should refer to, etc.) These issues, in turn, were fed back into the rest of the writing of this guidebook. This is a direct illustration of many of the core concepts discussed in this guidebook concerning both the strengths and inherent biases of any representation.

Appendix C: Canvas Lexicon

Lexicon Conventions

This lexicon defines terms used in the document. Within the main body of the document, these terms when first referenced in any given section appear in ***bold italic*** type. In later references within the same section they will generally appear in regular type, but may occasionally be repeated in ***bold italic*** type. Within the lexicon itself, the defined term is listed, for readability, in **bold regular** type on a single line.

Three types of terms are included in the lexicon:

- *General descriptive terms* for concepts essential to Canvas. We have tried to include only terms that are used with special meaning in the Canvas context. Some of these terms are used frequently in other disciplinary contexts (e.g., "informant" has typical conventions of usage in social scientific research) or general usage (e.g., "audience"). These terms are used in a specific technical sense in Canvas, i.e., to serve as an umbrella for several more specific terms, or to make key distinctions that would otherwise be difficult to convey.

  If you see a familiar term in ***bold italic*** in the main text of the document, make sure you are clear about the specific meaning it has in the Canvas context. For example, the term ***interaction*** is used here as a technical term for a person-to-person encounter between an investigator and an informant, in contrast to a ***study*** which involves and investigator and an inanimate source of information like a document. Terms introduced in the text may be in singular or plural form, depending on context, but entries in the lexicon are in singular.

- *Formal and informal terms.* In many cases a shortened or informal term is used extensively throughout the document (e.g., "session") in place of a longer phrase that disambiguates the intended meaning of the term (e.g., "knowledge acquisition session"). In these cases, we provide the main definition entry for the *more* formal term, using brackets in the term to indicate that part of the fuller and more formal phrase elided in common usage within the document (or, in some cases, variant phrases that might be used for the defined term). For example, in the entry "[knowledge acquisition] session", "knowledge acquisition" is bracketed to show that the term sometimes appears in the text as just "session". A cross reference is included in the lexicon under the less formal term (e.g., "session").

- *Non-Canvas terms.* A few entries may be included in the lexicon for terms used in general reuse literature that do not have special meaning in the Canvas context. These are included for clarity, context and completeness, not to suggest formal definitions for these terms.

In addition to the textual definitions of terms provided in this appendix, Appendix B includes formal notations of the semantic relations between some of the key terms in the lexicon. We encourage the reader to flip back and forth between these two appendices to see how the terms relate.

The approach to lexicon structure and organization presented here can, like Appendix B, be taken as one form of representation for knowledge acquisition. To be more precise, we have found that control of terminology is important in almost any representation used in KA. A lexicon, unlike a simple glossary or data dictionary, defines a set of terms within the context of a specific community of practice. The ODM method goes into considerable detail about the use of domain lexicon as a means of managing the knowledge acquisition process.

Lexicon Terms and Definitions

activity

A process or work task performed by a practitioner within a work setting. Activities can be observed by investigators as one form of knowledge acquisition. Note that within the SEP methodology this term has a more constrained and formal definition as part of scenarios. Within Canvas the main significance is that the activity serves as the effective knowledge source for sessions where there is no explicit interviewing of an informant or study of an existing artifact. See also practitioner, work setting, observation.

artifact

A workproduct developed within a work setting, that is used to provide information about the work setting in which it was created or used. An artifact is a type of knowledge source. See also workproduct, work setting, knowledge source.

audience

For any particular knowledge acquisition workproduct, its audience is anyone who can examine the workproduct, and thereby gain some of the knowledge that was the subject of the learning process that produced the workproduct. The audience can be drawn from any community. In some cases the audience community is the same as the knowledge source community or the investigator community. See also knowledge acquisition workproduct and learning.

Canvas

A knowledge acquisition approach that encompasses the idea that each participant (informant, investigator) in knowledge acquisition will be influenced by the ongoing knowledge acquisition effort. The name "Canvas" come from the image of weaving together "threads" corresponding to the lifecycle of each participant. Each thread is monitored and managed for the effect of changing biases throughout the lifecycle. In Canvas, the knowledge acquisition effort is viewed as a way to bridge the cultural gap between two (or more) communities of practice, by explicitly handling the variance as well as the commonality of view both within a single community and between communities. See also participant, knowledge acquisition, thread, knowledge acquisition effort, community of practice.

codification

The transfer of knowledge into a workproduct. Knowledge has been codified into a workproduct, if that knowledge can be learned through examination of the workproduct by a practitioner in the intended audience for the workproduct. See also workproduct.

collaboration

See collaborative knowledge creation.

collaborative knowledge creation

The process by which interaction between investigator and informant creates useful knowledge that was not available to either of them alone. A particular type of learning process resulting from a session with an informant in which the informant participates to some extent in the codification of the knowledge elicited. the investigator gains new knowledge that was not already held by the informant. Typically in such a situation, the informant learns also.

In contrast to reflection, the new knowledge contains some aspects of information unavailable to each participant; hence it could not have been produced by any single participant alone. See also learning, informant, investigator, reflection, knowledge transfer.

community of practice

A group of practitioners that share terminology, knowledge, and a set of behaviors and interests in a certain domain. See also work setting.

configuration

See knowledge transfer configuration.

derivative workproduct

A workproduct produced by interpreting some other workproduct, including another derivative workproduct. See workproduct, knowledge acquisition session, artifact.

dossier

A collection of all the knowledge acquisition workproducts created as part of a knowledge acquisition enterprise. Includes links to the artifacts and informants used as knowledge sources for the enterprise. See also knowledge acquisition workproduct, knowledge acquisition enterprise.

embedded [knowledge]

Knowledge not held in a form that can be easily articulated by any single member of a community of practice. The knowledge may exist in close alignment with social interactions. For example, an operating room team may have certain competencies in common that were developed over long periods of intensive work practice. The knowledge may be informal and not codified, passed on by direct contact and experiential training. The knowledge may also be embedded because it is part of the surrounding cultural context and hence not deemed as a characteristic aspect of the work setting itself, but is of significance to the investigator and/or target communities.

enterprise

See knowledge acquisition enterprise.

enterprise objective

See knowledge acquisition objective.

focus community

The community of practice in which the focus of interest is embedded. Practitioners from this community are selected as informants. Contrast to investigator community, target community. See also focus of interest, informant.

focus of interest

The topic in the knowledge source's work setting about which the investigator wishes to acquire information. The knowledge source should be knowledgeable about this topic. See topic, knowledge source, work setting.

informant

A person that serves the role of knowledge source in a knowledge acquisition session. Most typically an informant is a practitioner in the focus community. Knowledge acquisition specialists in the knowledge engineering field tend to use the term "domain expert" to describe the people that they interview in knowledge acquisition. We prefer a less specific term, since some people interviewed will neither consider themselves experts nor be considered such by their community. See also knowledge source, practitioner.

[knowledge acquisition] interaction

A session where knowledge is elicited through the interaction of two or more people; e.g., an interview or a facilitated group KA session. Contrast with a study, where people interact with artifacts. See also session, study.

investigator

A person who performs knowledge acquisition activities to acquire information from knowledge sources. Investigators are part of the KA team and have primary responsibility for the KA sessions in which knowledge is elicited. See also session, knowledge source, audience.

investigator community

The community of practice that is performing the knowledge acquisition. Contrast with focus community and target community. See also knowledge acquisition session.

KA

See knowledge acquisition.

KA plan

See knowledge acquisition plan.

knowledge acquisition

A special kind of knowledge creation process with the following properties:

- Knowledge is *elicited* from a knowledge source from one work setting by an investigator from another work setting.
- Some portion of the elicited knowledge is *codified* into a knowledge acquisition workproduct.
- The codified knowledge is transferred to an audience in a different community of practice.

Knowledge elicitation and codification activities take place in a knowledge acquisition session. They could take place in a single or in separate sessions.

The knowledge that has been effectively codified is that knowledge that can be learned by examining the resulting knowledge acquisition workproduct, by a person who was not present at the knowledge acquisition session, but who is in the intended audience for the workproduct. In any knowledge acquisition session, there is always more knowledge created than is effectively codified in the resulting workproduct. See also knowledge source, investigator, work setting, knowledge acquisition session, knowledge creation, codification.

knowledge acquisition effort

That portion of an overall project concerned with the systematic acquisition of knowledge. The effort might be part of an expert systems development project, a conventional systems development effort utilizing knowledge acquisition techniques, a domain engineering effort for the purposes of engineering reusable software assets, or for non-technical purposes such as the gathering of social scientific data. The effort encompasses the people (e.g., investigators, informants, sponsors, etc.), the processes (e.g., interviews, studies of artifacts), and the resulting knowledge acquisition workproducts.

On occasion, it is convenient to refer to the complete set of KA activities performed as part of an overall project as either the KA enterprise or the KA effort (depending on context). See also knowledge acquisition enterprise.

[knowledge acquisition] enterprise

Those aspects of a knowledge acquisition effort that are relatively constant throughout the effort (e.g., the objectives, intended audience, dossier infrastructure and resource pools). In this sense, in the enterprise level of KA planning is contrasted with phase, thread and session planning levels. Knowledge acquisition planning at the enterprise level is the first and foremost place where stakeholder issues are examined. Contrast to threads and sessions, many of which can be planned within the scope of a single knowledge acquisition effort.

On occasion, it is convenient to refer to the complete set of KA activities performed as part of an overall project as either the KA enterprise or the KA effort (depending on context). See also knowledge acquisition effort, stakeholder.

knowledge acquisition [KA] mode

A characteristic pattern or "topology" for a knowledge transfer configuration. One mode involves transfer from a focus community back to the same community, through the mediation of an investigator community. In this mode, the codification of the knowledge in a new form and the resulting transformation of the knowledge transfer mechanisms within the community, is the primary goal of the KA activity. A second mode involves transfer of knowledge to the investigator community itself. This is the typical mode for an academic research project, where knowledge about the focus community is of direct strategic value for the goals of the investigator community. A final mode, and the one most relevant for this document's primary audience, involves distinct focus, investigator and target communities, where the investigator community is acting as a "bridge" or, literally, transfer agent between the other two communities.

[knowledge acquisition] objective

A specific goal for a knowledge acquisition activity. Objectives are set for the knowledge acquisition enterprise as a whole, which in turn guides determination of more specific objectives for particular knowledge acquisition threads and sessions.

- **Enterprise objectives** include transfer of a specific scope of knowledge across specified communities of practice that enable certain new work to be accomplished. (This is modeled after the learning objectives used in training development, which seek to measure effective learning in terms of what learners will be able to do that they were not able to do previously.)
- **Thread objectives** involve the desired end result of the series of sessions that constitute the thread for the KA element in question. For example, objectives for an investigator's thread might be to achieve broad knowledge of a subset of topics in the overall domain, and to have achieved a given new level of experience in facilitating KA group sessions.

- **Session objectives** include the topics of focus for the session, the knowledge types to be acquired, desired level of detail, etc.

See also topic of focus, knowledge type, knowledge acquisition session, knowledge acquisition enterprise. knowledge transfer configuration.

knowledge acquisition plan

A centralized plan created to guide a KA enterprise. The plan covers enterprise, phase, thread and session-level planning, and is adaptive to allow for ongoing re-planning based on new information. See also knowledge acquisition objectives, knowledge acquisition enterprise.

[knowledge acquisition] session

A knowledge acquisition session is an event that involves at least one investigator and at least one knowledge source, and generally produces a knowledge acquisition workproduct. The fundamental types of sessions are studies of artifacts and interactions with informants. The audience for the workproduct created may be within the informant's setting, the investigator's setting or a separate target setting. See also investigator, knowledge acquisition workproduct, artifact, informant, work setting, audience.

[knowledge acquisition] workproduct

A particular workproduct that was created in the work setting of knowledge acquisition, that is, the result of a knowledge acquisition session. The knowledge acquisition workproducts are retained in the dossier. Knowledge acquisition workproducts also have an audience, which is a community of practice who can be expected to be able to comprehend it. See also workproduct, knowledge acquisition session, dossier, audience.

knowledge creation

The most general term for a process that results in new knowledge being created. Individual learning, formal teaching, documenting of process, research, and knowledge acquisition (the focus of this document) are all forms of knowledge creation. See knowledge acquisition.

knowledge engineer

In the expert systems context, someone familiar with particular knowledge representation systems and knowledge acquisition techniques. A knowledge engineer thus performs the tasks of an investigator utilizing more than knowledge of the domain of interest; i.e., they are not only informants, even though they may have considerable domain knowledge, or will acquire it over time. More importantly, a knowledge engineer is committed to developing knowledge acquisition and knowledge representation skills that are transferable across domains.

In the Canvas context, therefore, not all investigators are knowledge engineers. We also extend the notion to apply to the skills of mastering a repertoire of different representations, with knowledge of the uses, inherent biases and appropriate elicitation techniques for each. In particular, mere familiarity of one knowledge representation system, without knowledge of the contextual aspects of its use in KA, would not render one a knowledge engineer in this broader sense.

knowledge source

Anything that can provide information that is embedded in some work settings. Human knowledge sources are called informants, while inanimate knowledge sources are called artifacts. See informant, artifact.

[knowledge] transfer

At the enterprise level, that aspect of knowledge acquisition which involves the transfer of knowledge from one community of practice to another by introducing a new codified form for the knowledge. At the individual level, a particular type of learning process resulting from a session with a learner and a knowledge source, in which the learner gains knowledge already held by the knowledge source. See also collaboration, learning.

knowledge transfer configuration

A schematic diagram that indicates a knowledge acquisition scenario in terms of named communities of practice that are assigned the respective roles of focus, investigator and target communities for a specific knowledge acquisition objective for a KA enterprise. A knowledge transfer configuration (appropriately annotated) is the primary documentation for a specific KA enterprise objective. See also knowledge acquisition objective, knowledge acquisition transfer mode.

[knowledge] transfer mode

A commonly recurring pattern for the transfer of knowledge between focus, investigator and target communities. The modes are useful because they highlight characteristic stakeholder issues that will arise in given configurations. They also provide the base patterns for the knowledge transfer configurations used in determining KA enterprise objectives. See also knowledge transfer configuration.

knowledge type

A broad categorization of knowledge that distinguishes static vs. dynamic knowledge, procedural vs. definitional knowledge, etc. There are many diverse systems for classifying types of knowledge and no particular system is critical to the integrity of the Canvas approach, in which knowledge types serve primarily as an intermediary level between specific topics or domains of knowledge and strategies for representation. Knowledge types are useful in two ways: they allow for correlating multiple specific topics with representations effective for codifying knowledge in those areas; conversely, each topic or knowledge area can be viewed from the aspect of multiple knowledge types. So, for example, definitional or categorical knowledge exists in most domains or topic areas, and can usually be represented well with taxonomic representations. Knowledge about individual skills or learned competencies involves procedural knowledge for which process descriptions or task diagrams may be useful. See also representation, topic.

learning

Any process by which a person gains some knowledge. See also knowledge transfer, collaboration and reflection, which are types of learning.

notation

The mode of representation of the information in a workproduct. In general, the notation could be any of natural language, diagrams, charts, tables, mathematical formula or the like. In Canvas, we usually consider workproducts within representations particular to knowledge acquisition. The notation in which a workproduct is represented constrains its possible content, and influences who its audience can be. For example, programming language representations are not usually accessible to medical personnel. See also representation, workproduct.

objective

See knowledge acquisition objective.

observation

An investigator eliciting knowledge by being directly present in a work setting and passively observing work activities being performed by practitioners. Whether the observed practitioners are informants or not is a grey area that depends on their degree of awareness of the presence of the investigator, the extent to which their activities are altered as a result of the observation, etc. Also, if a passive recording medium such as audio or video tape is used, one must consider the session in which this knowledge acquisition workproduct is reviewed by investigators as part of the chain of interpretation for the event recorded. See also informant, knowledge acquisition workproduct.

practice

A recurring process or set of activities that takes place in a given work practice setting. The activities are considered by the members of the community of practice to be part of the work that goes on in that setting. See also setting, community of practice.

practitioner

A member of a community of practice. Practitioners are always humans. See also community of practice, informant, artifact.

project context

The broader project within which a KA enterprise is performed. Typically, given the intended audience for this document, this would be a technology development effort where some degree of systematic KA is deemed of value.

reflection

A particular type of learning in which the learner gains knowledge without resorting to a knowledge source. Reflection is, therefore, not dependent on a knowledge acquisition session. See also learning, transfer, collaboration, knowledge acquisition session.

repertoire [of representations]

In the Canvas context, a repertoire of representations is the set of representations used by a specific community of practice. The repertoire includes both the set of representations and notations and the "theory of use" that allows them to be used together or in complementary ways. Skilled knowledge engineers should develop a broad repertoire of representations to allow them to interact with many different practitioner communities.

representation

A description of some topic, produced during a knowledge acquisition session, and recorded in a workproduct. A representation includes a notation and associated semantics. A given community of practice generally has a set or repertoire of representations in use. Knowledge engineers work with repertoires of representations as part of their KA skill-set. See also workproduct, notation, repertoire.

session

See knowledge acquisition session.

session objectives

See knowledge acquisition objectives,

setting

See work setting.

stakeholder

A person, group, or organization that has interests and objectives relative to the knowledge acquisition effort. Stakeholders have diverse viewpoints, experience, and terminology. See also knowledge acquisition effort.

study

An knowledge acquisition session where an investigator examines an artifact to elicit knowledge about a work setting. Contrasted with interaction. See investigator, artifact, interaction, session.

systematic knowledge acquisition

A systematic approach to knowledge acquisition provides a repeatable procedure for making key decisions in planning and performing knowledge acquisition, and for recording the results of knowledge acquisition activities in such a way that essential contextual information about the data acquired is preserved.

target community

In a knowledge acquisition project involving three distinct communities, the community that is the intended audience of the generated workproducts. A project in which the audience is either the focus or investigator community will not have a separate target community. When speaking generally, we refer to the audience community as the target community, whether it is the same as the focus or investigator community, or a third, separate community. Contrast to focus community and investigator community. See also audience, workproduct.

thread

The lifecycle of an informant, artifact, or investigator. The word "thread" is intended to convey the image of linearity of a lifecycle, interacting in a two-dimensional "canvas". Each entity in the knowledge acquisition project (informant, artifact, or investigator) follows a thread through the fabric of the overall project. See also informant, artifact, investigator.

thread objective

See knowledge acquisition objective.

topic [of focus]

A specific criterion for selecting and focusing information gathered in a knowledge acquisition session; part of the knowledge acquisition objectives for the session. See also knowledge acquisition objectives.

transfer

See knowledge transfer.

transfer mode

See knowledge transfer mode.

work practice setting

See work setting.

workproduct

The result of an activity in any work setting. In order to be a workproduct, the result must be in a form that can be accessed by someone who was not involved in the activity. Thus reports, tapes, documents, programs and diagrams are all workproducts, but new ideas that have not been captured on paper are not workproducts. See also work setting, knowledge acquisition workproduct.

[work] setting

An environment where people interact with each other and perform processes. Settings imply a certain stability in that the same people work together on a routine basis. This set of people forms a community of practice. See also community of practice.

Bibliography

In the following bibliography, documents which pertain most directly to ODM and its application are denoted with an asterisk (*) Documents which pertain to SEP and its applications are denoted with a double asterisk (**).

[1] * Army STARS Demonstration Project. Domain Engineering Guidebook. U.S. Army CECOM Software Engineering Directorate, Ft. Monmouth NJ, 1995.

[2] Bailey, K. Typologies and Taxonomies: An Introduction to Classification Techniques. Sage University Paper, Series: Quantitative Applications in the Social Sciences, Series/Number 07-102, Sage Publications, 1994.

[3] Boehm, B., and W. Scherlis. "Megaprogramming." In *Proceedings of the DARPA Software Technology Conference*, Arlington VA, April 1992.

[4] Bojie, D. "The Storytelling Organization: A Study of Story Performance in a Office-supply Firm." *Administrative Science Quarterly*, Vol. 36, 1991, pp.106-26.

[5] Brown, J. S. "Research that Reinvents the Corporation." *Harvard Business Review*, March-April 1991.

[6] Buchanan, B.G., G.L. Sutherland, and E. A. Feigenbaum. "Rediscovering Some Problems of Artificial Intelligence in the Context of Organic Chemistry." in B. Beltzer and D. Michie (eds) Machine Intelligence vol. 5, Edinburgh University Press, pp. 253-280, 1969.

[7] Clancey, W. J. From Guidon to Neomycin and Heracles in Twenty Short Lessons. AI Magazine, 1986.

[8] * Collins, P. "Toward a Reusable Domain Analysis." In *Proceedings of the 4th Annual Workshop on Software Reuse*, Herndon VA, November 1991.

[9] Creps, R., M. J. Davis, M. Simos, et al. "Using a Conceptual Framework for Reuse Processes as a Basis for Reuse Planning." In *Proceedings of the 7th Annual Software Technology Conference*, Salt Lake City UT, April 1995.

[10] Duda, J., J. Gaschnig, and P. Hart, "Model Design in the Prospector Consultant System for Mineral Exploration," in D. Michie 9ed) Expert Systems in the Microelectronic Age, Edinburgh University Press, Edinburgh, pp. 153-167, 1979.

[11] Foreman, J. "STARS Mission." In *Proceedings of the DARPA Software Technology Conference*, Arlington VA, April 1992.

[12] Freeman, M. W.,L. Hirschman, D. P. McKay, F. L. Miller, and D. P. Sidhu. Logic Programming Applied to Knowledge-Based Systems, Modelling, and Simulation. In *Proceedings of the Conference on Artificial Intelligence*, pages 177-193, April 1983.

[13] Frey, D. "Something's Got to Give." New York Times Magazine, March 24, 1996.

[14] Friedrichs, R. *A Sociology of Sociology.* The Free Press/Collier Macmillan, 1970.

[15] **Haddock, G., and K. Harbison. "From Scenarios to Domain Models; Processes and Representations," in *Proceedings of Knowledge-based Artificial Intelligence Systems in Aerospace and Industry. SPIE: International Society for Optical Engineering*, April 1994.

[16] Henninger, S. "Accelerating the Successful Reuse of Problem Solving Knowledge Through the Domain Lifecycle." In *Proceedings of the Fourth International Conference on Software Reuse*, IEEE Computer Society Press.

[17] Johnson, J. *Selecting Ethnographic Informants.* Sage University Paper, Series: Qualitative Research Methods, Volume 22, Sage Publications, 1990.

[18] Jordan, B. "Cosmopolitical Obstetrics: Some Insights from the Training of Traditional Midwives," *Soc. Sci. Med.* Vol. 28, No. 9, pp. 925-944, 1989.

[19] * Kelly, T.P., W. Lam, and B.R. Whittle. "Diary of a domain analyst: a domain analysis case-study from avionics." In *Proceedings of IFIP Working Groups 8.1/13.2 Conference on Domain Knowledge for Interactive System Design*, Geneva, May 1996.

[20] Keough, D.M. "Creating the Learning Organization: An Interview with Peter Senge," *The McKinsey Quarterly*, 1992.

[21] Klingler, C. "A Practical Approach to Process Definition." In *Proceedings of the 7th Annual Software Technology Conference*, Salt Lake City UT, April 1995.

[22] Linde, C. "Reflections on Workplace Learning." Working Paper, Institute for Research on Learning, 2550 Hanover St., Palo Alto CA, 1993.

[23] Mancuso, J.C., and M.L.C. Shaw. *Cognition and Personal Structure: Computer Access and Analysis.* Praeger Press, New York NY, 1988.

[24] Marca, D., and C. McGowan. *SADT, Structured Analysis and Design Technique.* McGraw-Hill, New York NY, 1988.

[25] ** Mettala, E., K. Harbison and S. Hufnagel. Scenario-based Engineering Process for Reconnaissance, Surveillance and Target Acquisition, in Proceedings of the ARPA Image Understanding Workshop, Morgan-Kaufman, November 1994.

[26] Meyer, M.A., J.M. Booker, J.M. Bradshaw. "A flexible six-step program for defining and handling bias in knowledge acquisition." in Wielinga, Boose, Gaines, Schreiber, and van Someren (eds), *Current Trends in Knowledge Acquisition.* IOS Press, Amsterdam, 1990.

[27] Moore, G. *Crossing the Chasm: Marketing and Selling Technology Products to Mainstream Customers.* Harper Business, New York NY, 1991.

[28] Noblit, G., R. Hare. *Meta-Ethnography: Synthesizing Qualitative Studies.* Sage University Paper, Series: Qualitative Research Methods, Volume 11, Sage Publications, 1988.

[29] O'Connor. B. *Healing Traditions: Vernacular Health Traditions and their Implications for Clinicians.* University of Pennsylvania Press, Philadelphia, 1994.

[30] Schein, E. *The Clinical Perspective in Fieldwork*. Qualitative Research Methods Series 5, Sage, Newbury Park CA, 1987.

[31] Schon, D. *The Reflective Practitioner: How Professionals Think in Action*. Basic Books, new York NY, 1983.

[32] Schwartz, P. *The Art of the Long View, Planning for the Future in an Uncertain World*. Doubleday/Currency, New York NY, 1991.

[33] Schwartzman, H. *Ethnography in Organizations*. Qualitative Research Methods Series 27, Sage, Newbury Park CA, 1993.

[34] Senge, P. *The Fifth Discipline*. Doubleday/Currency, New York NY, 1990.

[35] Senge, P., et al. *The Fifth Discipline Fieldbook: Strategies and Tools for Building a Learning Organization*. Doubleday/Currency, New York NY, 1994.

[36] Shaw, M. L. G. and C. McKnight, *Think Again: Personal Decision Making and Problem Solving*. Prentice Hall, Hemel Hempstead, 1981.

[37] * Simos, M. "The Growing of an Organon: A Hybrid Knowledge-Based Technology and Methodology for Software Reuse." In *Domain Analysis and Software Systems Modeling*, R. Prieto-Diaz and G. Arango, ed., IEEE Computer Society Press, 1991.

[38] * Simos, M., "Juggling in Free Fall: Uncertainty Management Aspects of Domain Analysis Methods." In *Proceedings of the 5th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*, Springer-Verlag, July 1994.

[39] * Simos, M., "Organization Domain Modeling (ODM): Formalizing the Core Domain Modeling Life Cycle." *SIGSOFT Software Engineering Notes*, Special Issue on the 1995 Symposium on Software Reusability, Aug 1995.

[40] * Simos, M., "Domain Modeling Techniques for Representing Commonality and Variability: Towards a Comparative Framework." *Proceedings of the 7th Annual Workshop on Software Reuse*, St. Charles, IL, August 1995.

[41] Softech, Inc. "Integrated Computer-Aided Manufacturing (ICAM) Architecture Part II. Volume IV - Function Modeling Manual ($IDEF_0$)." Technical Report AFWAL-TR-81-4023 Volume IV, Materials Laboratory (AFWAL/MLTC), AF Wright Aeronautical Laboratories (AFSC), Wright-Paterson AFB, Dayton OH, June 1981.

[42] Spradley, J. *The Ethnographic Interview*. Holt, Rinehart, and Winston, New York NY, 1979.

[43] Spradley, J. *Ethnographic Observation*. Holt, Rinehart, and Winston, New York NY, 1979.

[44] Sridharan, N. S., Guest editorial for Special Issue on Applications to Science and Medicine, Artificial Intelligence, vol 11, pp 1-4, 1978.

[45] * STARS. Army STARS Demonstration Project Experience Report. STARS Technical Report STARS-VC-A011R/003/02, STARS Technology Center, Arlington VA, April 1996.

[46] STARS. STARS Conceptual Framework for Reuse Processes (CFRP), Vol. I: Definition, Version 3.0. Unisys STARS Technical Report STARS-VC-A018/001/00, STARS Technology Center, Arlington VA, October 1993.

[47] STARS. STARS Conceptual Framework for Reuse Processes (CFRP), Vol. II: Application, Version 1.0. Unisys STARS Technical Report STARS-VC-A018/002/00, STARS Technology Center, Arlington VA, September 1993.

[48] STARS. Learning and Inquiry Based Reuse Adoption (LIBRA): A Field Guide to Reuse Adoption through Organizational Learning, Version 1.1. Loral Defense Systems-East STARS Technical Report STARS-PA33-AG01/001/02, STARS Technology Center, Arlington VA, February 1996.

[49] * STARS. Organization Domain Modeling (ODM) Guidebook, Version 2.0. STARS Technical Report STARS-VC-A025/001/00, Lockheed Martin Tactical Defense Systems, Manassas VA, June 1996.

[50] STARS. OpenRLF Technical Paper. STARS Technical Report STARS-PA31-AE08/001/00, Lockheed Martin Tactical Defense Systems, Manassas VA, June 1996.

[51] STARS. Reuse Strategy Model: Planning Aid for Reuse-Based Projects. Boeing STARS Technical Report D613-55159, STARS Technology Center, Arlington VA, July 1993

[52] Trist, E. "Referent Organizations and the Development of Interorganizational Domains." *Human Relations*, Vol. 36, No. 13, pp. 269-84, 1983.

[53] * Unisys. Reusability Library Framework AdaKNET and AdaTAU Design Report. PAO D4705-CV-880601-1, Unisys Defense Systems, System Development Group, Paoli PA, 1988.

[54] Unisys. RLF Modeler's Manual, version 4.2, Unisys Defense Systems, System Development Group, 1993.

[55] ** Wang, W., S. Hufnagel, P. Hsia, S. M. Yang, "Scenario Driven Requirements Analysis Method." *Proceedings of the Second International Conference on Systems Integration*, June 15-18 1992, Morristown, NJ, IEEE.

[56] Weisbord, M. *Discovering Common Ground*. Berrett-Koehler, San Francisco CA, 1992.

[57] Weizenbaum, J. *Computer Power and Human Reason*, W. H. Freeman and Co., San Francisco, 1976.

[58] Welbank, M. British Telecom Report on Knowledge Acquisition, British Telecom, London 1983.

[59] Winograd, T., and F. Flores. *Understanding Computers and Cognition: A New Foundation for Design*. Addison-Wesley, Reading MA, 1987.

[60] Zuboff, S. *In the Age of the Smart Machine*. Basic Books, New York NY, 1988.